BMJ Open

Two-year effectiveness of a stepped-care depression prevention intervention and predictors of incident depression in primary care patients with diabetes type 2 and/or coronary heart disease and subthreshold depression

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020412
Article Type:	Research
Date Submitted by the Author:	01-Nov-2017
Complete List of Authors:	Pols, Alide; VU University, Department of Health Sciences Adriaanse, Marcel; Institute of Health Sciences, Vrije Universiteit Amsterdam Tulder, Maurits; University of Amsterdam, Health Sciences Heymans, Martijn; VU University Medical Center, Department of Epidemiology and Biostatistics; VU University, Department of Health Sciences Section Methodology and Applied Biostatistics Faculty of Earth and Life Sciences Bosmans, J; VU University Amsterdam van Dijk, Susan; Vrije Universiteit Amsterdam, Department of Health Sciences Van Marwijk, H; The University of Manchester, Centre for Primary Care
Keywords:	Depression & mood disorders < PSYCHIATRY, DIABETES & ENDOCRINOLOGY, Coronary heart disease < CARDIOLOGY, PREVENTIVE MEDICINE

Two-year effectiveness of a stepped-care depression prevention intervention and predictors of incident depression in primary care patients with diabetes type 2 and/or coronary heart disease and subthreshold depression

Alide D Pols*[1,2], Marcel C Adriaanse[1], Maurits W van Tulder[1], Martijn W Heymans[3], Judith E Bosmans[1], Susan E van Dijk[1], Harm W J van Marwijk[2,4]

[1] Department of Health Sciences, Amsterdam Public Health research institute, Faculty of Science, Vrije Universiteit, Amsterdam, The Netherlands

[2] Department of General Practice & Elderly Care Medicine, Amsterdam Public Health research institute, VU University Medical Centre, Amsterdam, The Netherlands

[3] Department of Clinical Epidemiology and Biostatistics VU Medical Centre Amsterdam, The Netherlands

[4] CLAHRC Greater Manchester and NIHR School for Primary Care Research, the University of Manchester, United Kingdom

* Corresponding author: a.d.pols@vu.nl

Keywords: major depressive disorder, subthreshold depression, diabetes mellitus type 2, coronary heart disease, effectiveness, stepped care, prediction model

ABSTRACT

Introduction Major depressive disorders (MDD), diabetes mellitus type 2 (DM2) and coronary heart disease (CHD) are leading contributors to the global burden of disease and often co-occur.

Objectives To evaluate the two-year effectiveness of a stepped-care intervention to prevent MDD compared to usual care and to develop a prediction model for incident depression in DM2 and/or CHD patients with subthreshold depression.

Methods Data of 236 Dutch primary care DM2/CHD patients with subthreshold depression (Patient Health Questionnaire 9 (PHQ-9) score ≥6, no current MDD according to the Mini International Neuropsychiatric Interview (DSM-IV criteria), who participated in the Step-Dep trial were used. A PHQ-9 score of ≥10 at minimally one measurement during follow-up (at 3, 6, 9, 12 and 24 months) was used to determine the cumulative incidence of MDD. Potential demographic and psychological predictors were measured at baseline via web-based self-reported questionnaires and evaluated using a multivariable logistic regression model. Model performance was assessed with the Hosmer–Lemeshow test, Nagelkerke's R^2 explained variance and Area Under the Receiver Operating Characteristic curve (AUC). Bootstrapping techniques were used to internally validate our model.

Results 192 patients (81%) were available at two-year follow-up. The cumulative incidence of MDD was 97/192 (51%). There was no statistically significant overall treatment effect over 24 months of the intervention (OR 1.37; 95% CI 0.52; 3.55). Baseline levels of anxiety, depression, the presence of >3 chronic diseases and stressful life-events predicted the incidence of MDD (AUC 0.80 interquartile range (IQR) 0.79-0.80; Nagelkerke's R^2 0.34 IQR 0.33-0.36).

Conclusion A model with four factors predicted depression incidence during two-year follow-up in patients with DM2/CHD accurately, based on the AUC. The Step-Dep intervention did not influence the incidence of MDD. Future depression prevention programs should target patients with these four predictors present, and aim to reduce both anxiety and depressive symptoms.

TRIAL REGISTRATION NUMBER

Dutch Trial Register NTR3715 http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=3715

STRENGTHS AND LIMITATIONS OF THIS STUDY

- This study provides a prediction model of incident MDD in DM2 and/or CHD patients with subthreshold depression, which could assist in its detection, enable healthcare providers to facilitate targeting indicated prevention to highest risk patients
- Only predictors that are readily available or easily obtained in practice were used in the multivariable model, which enhances the practical use of the model
- This study had a relatively long follow-up and outcomes were frequently measured, whereas drop-out rates were relatively low and missing values imputed
- The relatively small study population might have caused over-optimism of the prediction model, but an internal validation procedure with bootstrapping techniques showed that this risk was minor
- Data were derived from a RCT, but statistically non-significant intervention effects for incident MDD at both 12- and 24-months follow-up justify using the Step-Dep population as a cohort

INTRODUCTION

Depression is a major and increasing contributor to the global burden of disease[1], whereas coronary heart disease (CHD) and diabetes mellitus type 2 (DM2) rank among the leading causes of morbidity and mortality worldwide[2]. Comorbid depression in patients with DM2 and/or CHD is common[3,4] and has detrimental effects on self-care and medication adherence[5,6], quality of life[7], health status and increases healthcare costs[8,9] and mortality[10,11]. Despite its negative impact, many cases of depression go unrecognized in primary care[12], especially in patients with chronic diseases like DM2 and/or CHD[13]. Additionally, about one-third of those recognized and treated does not respond to current approaches, and over half of those who experience a first onset of a major depressive episode will experience one or more recurrences[14].

Given the significant burden of disease, the poor recognition of and limited effect of current treatment options for depression, it would be of great value if incident cases could be averted by early detection and preventive treatment of patients at risk ('indicated prevention'). Meta-analyses have shown that preventive psychological interventions can overall reduce the incidence of MDD in comparison to control groups[15,16]. Offering preventive psychological interventions in a stepped-care format could be an efficient approach, as patients start with minimally intensive evidence-based treatments and only those who do not improve adequately, step up to a treatment of higher intensity[17]. Recently, we conducted a randomized controlled trial in which we evaluated whether a pragmatic nurse-led stepped-care program was effective in reducing the incidence of MDD at 12-months of follow-up in comparison with usual care among patients with DM2 and/or CHD and subthreshold depression (Step-Dep)[18]. Subthreshold depression entails clinically relevant depressive symptoms without fulfilling the criteria for MDD and is a known important risk factor for depression[15,19]. We demonstrated that the Step-Dep intervention was not superior to usual care and the overall cumulative incidence of MDD was lower than expected after one year [20]. However, it may be possible that the follow-up period was too restricted to demonstrate the potential health benefits of the stepped-care program over usual care, or the presence of subthreshold depression alone posed a lower than expected prior risk of MDD in our DM2 and/or CHD population.

Identifying additional major risk factors of incident depression in patients with DM2 and/or CHD might facilitate targeting indicated prevention to patients with highest risk, but also potentially aid in its detection. In patients with DM2, several longitudinal studies have been conducted to determine risk factors for comorbid incident depression. However, these studies have rendered heterogeneous results, due to small patient samples (<80 at follow-up)[21,22], analyses of single factors only[23,24], the use of mixed samples of type 1 diabetes and DM2[25], patients with either no MDD at baseline[23,26] or both with and without depression at baseline[22,24,25,27], and differences across community[23,24], primary care[25,27] and secondary care settings[22,26]. In patients with CHD, the only available longitudinal data are derived from studies in patients with acute coronary syndrome followed-up after hospital discharge[28–32]. Predictors that were repeatedly identified in DM2 or CHD studies were: depression severity at baseline[21,22,25,28,31,32], history of depression[25,27,29], female sex[24,27,29,31] and baseline anxiety levels[21,30,31]. However, data of patients with both DM2 and CHD, non-acute CHD or within primary care settings are scarce. The goal of the present study was twofold: (1) to evaluate the two-year effectiveness of a nurse-led stepped-care intervention to prevent MDD as compared to usual care (Step-Dep); and to (2) develop a prediction model for incident depression during two-year follow-up in primary care patients with DM2 and/or CHD and subthreshold depression.

METHODS

Design

Data of the Step-Dep cluster randomized controlled trial were used. Step-Dep was conducted in 27 general practitioner (GP) practices in three regions in the Netherlands (Amsterdam, Leiden, Twente), between January 2013 and November 2016, including recruitment and two years of follow-up. Details on the methods and design of the Step-Dep study have been published elsewhere[18].

Patients

Included patients were aged 18 years or more who had an International Classification of Primary Care (ICPC) diagnosis of DM2 and/or CHD and had subthreshold depression identified by screening. Patients with a Patient Health Questionnaire 9 (PHQ-9; range 0-27 with higher scores indicating more severe depressive symptoms) score of six or higher[33,34], and no major depressive disorder according to the Mini International Neuropsychiatric Interview (MINI)[35,36], were considered to have subthreshold depression. Exclusion criteria were cognitive impairment, psychotic illnesses, a terminal illness, the use of anti-depressant medication, a history of suicide attempt(s), loss of significant other in the past six months, visual impairment, current pregnancy, bipolar disorder, borderline personality disorder or any difficulties completing written questionnaires or visiting the primary care center. A total of 236 patients gave informed consent to participate.

Outcome measure

The outcome measure used was an incident depression (yes/no) defined as a PHQ-9 score of ≥10 at minimally one moment during follow-up (measured at 3, 6, 9, 12 and 24 months after baseline). The PHQ-9 is a widely used and validated instrument that performs well in patients with chronic medical illnesses both as dichotomous diagnosis of major and minor depression and a continuous severity score[34,37]. A cut-off of ≥10 has been shown to be the optimum cutoff for major depression[38], also in this patient group [39]. PHQ-9 was self-reported with web-based or written questionnaires. When these web-based or written questionnaires were not completed, the PHQ-9 was administered by telephone by trained research-assistants, blinded to randomization status.

Potential predictors

The selection of the potential predictors was based on a thorough literature search. Predictors of incident depression that were identified in multiple studies in patients with DM2 or CHD and are routinely available or easily obtained in daily GP practice were used. Additionally, we chose the presence of multiple chronic diseases[24] and stressful life-events[28] although they were identified in single studies only, as these were also indicated as causes of depression by patients and practice nurses in semi-structured interviews as part of the process evaluation of Step-Dep[40], and age[23].

Apart from GP information system derived data on *sex, age* and *ICPC diagnosis of DM2 and/or CHD*, demographics and psychological factors were measured at baseline via web based (or written if preferred) self-reported questionnaires. To take possible effects of the intervention into account, we included randomization status in the selection models as well. Patients in the intervention arm were offered a stepped care prevention program, and patients in the control arm received care as usual during one year. The stepped care intervention consisted of four sequential but flexible treatment steps, each lasting three months; 1) watchful waiting, 2) guided self-help, 3) problem solving treatment and 4) referral to a general practitioner. After each step, patients with a persisting PHQ-9 score of six or more were offered the next treatment step of the intervention. *Baseline depression levels* were measured with the PHQ-9[33,34]. *Baseline anxiety levels* were measured with the Hospital Anxiety and Depression Scale Anxiety (HADS-A; range 0-21 with higher scores indicating more severe anxiety)[41]. *History of depression* and s*tressful life-events were* self-reported using a subset of the Diagnostic Interview Schedule (DIS)[42]. *Number of co-morbid chronic illnesses* was measured using the self-reported Dutch Questionnaire Chronic Illnesses[43]. This was dichotomized using the median in our sample: three or less versus more than three chronic diseases.

Statistical analyses

The two-year effectiveness of the intervention on the primary and secondary outcomes was analyzed according to the intention to treat principle. Generalised Estimating Equations (GEE) were used for binary outcome variables, and linear mixed models for longitudinal data were used for continuous outcome variables[44]. For each outcome an overall effect over time and separate effects at different time points were estimated by taking time into

account as a categorical variable (with five categories: 0-3 months, 3-6 months, 6-9 months, 9-12 months and 12-24 months of follow-up)[45,46]. The main analyses consisted of fully corrected models that were corrected for baseline values of the respective outcome and additionally included the covariates gender[47], age[48], and any other possible confounding variable on which the treatment groups differed at baseline (marital status, employment status, level of education, co-existence of DM2 and CHD, alcohol use, number of depressive episodes in history and age of onset of depression), based on absolute baseline differences judged by the researchers rather than statistical testing[49]. For these analyses, STATA version 14 was used.

Missing data were imputed using multiple imputation according to the Multivariate Imputation by Chained Equations (MICE) algorithm[50] in SPSS version 23. For the imputations, missing at random (MAR) was assumed. Variables that were associated with missing data and variables that were associated with the outcome, were identified and included in the imputation model. Also, all variables in the analysis model (potential predictors and outcome) were included. The number of imputed datasets was 25 based on the proportion of cases with incomplete measurements; 24%. The subsequent analyses were performed on pooled data according to Rubin's rules[51].

Prediction model

We created a multivariable logistic regression model in SPSS 23 from the baseline variables estimating the probability of having at least one major depression (PHQ ≥10) during the two-year assessment. To calculate the number of potential predictors for developing the prediction model, we used the criterion of 10 events per variable. Continuous variables were checked for linearity with the outcome using spline regression curves and linearity was confirmed. All variables were entered into the logistic model and tested for statistical significance in the presence of the total set of predictors. Individually, the least significant predictor (P-value>0.157, as recommended in the TRIPOD statement, [52], Wald statistic) was removed, and the model was refit (backward selection). Randomization status was maintained in the model. This was repeated until we reached a statistical model that only included statistically significant predictors. This was repeated with p-values of 0.05. We also

compared the results with complete case analysis (CCA), i.e., all patients with missing data were excluded from the analyses.

We checked the performance of the model with regard to the goodness of fit (Hosmer–Lemeshow test), the explained variation and the discriminative ability of the model. The Nagelkerke's R^2 explained variation is the extent to which the outcome can be predicted by the predictors in the model in current datasets. The discriminative ability is reflected by the area under the receiver operating characteristics curve (AUC). Bootstrapping techniques were used to internally validate our model, i.e., to simulate the performance with respect to the explained variance and the AUC in comparable patient datasets[53]. After that, we calculated the linear predictor of the bootstrapped model with an adjusted intercept and regression coefficients corrected for the shrinkage factor. Performance measures were assessed in each imputed dataset and results were summarized using median values [54]. All analyses were done with SPSS version 23.0 and R software.

RESULTS

Participants

The baseline characteristics of the study population are presented in Table 1. Of the 236 patients included in Step-Dep, 192 patients (81%) completed two years of follow-up. A flowchart of participants through the first 12 months of the Step-Dep study has been published elsewhere[20]. At 24 months of follow up 18 additional patients dropped out (two for unknown motives, seven due to time considerations, four were deceased, three too frail, two unable to contact). We compared the baseline characteristics of patients with missing data to those without. Patients with missing data were more often living alone (61% vs 41%), but no other differences between these groups were found.

There was no statistically significant overall treatment effect over 24 months of the intervention (OR 1.37; 95% CI 0.52 to 3.55), nor at any of the time-points. There were no significant differences in PHQ-9 scores between the study groups at any time-point and the course of PHQ-9 scores over time did not differ significantly between the groups. Results are shown in Table 2. The statistically non-significant intervention effects for incident MDD at both 12-months[20] and 24-months of follow-up justify using the Step-Dep population as a cohort.

Prediction model

The cumulative incidence during two-year follow-up was 97/192 (51%). The multivariable models using p=0.05 and p=0.157[52] were identical. The final model consisted of four predictors: level of anxiety, level of depression, presence of more than three chronic diseases and having suffered a stressful life-event in the past year. This model performed well (Hosmer–Lemeshow test p=0.12 and median of pooled Nagelkerke's R^2 explained variance 0.34 interquartile range (IQR) 0.33-0.36) with good discriminative properties (median of the pooled AUC 0.80 IQR 0.79-0.80). In a CCA with p=0.05, the same predictors remained. In a CCA using p=0.157 [52], the categorical variable DM2/CHD/both also remained.

The risk of an incident MDD during two years of follow-up more than doubled when either more than three chronic diseases were present or a patient had suffered a stressful life-event in the past year. Both higher

depression and anxiety levels at baseline increased the risk of MDD with each incremental point on the PHQ-9 of HADS scales respectively. One point higher on the PHQ-9 at baseline, resulted in a 1.37 higher risk of developing MDD during two years, compared to 1.13 for increasing anxiety levels. With regard to the internal validation of the model, the calibration slope (or shrinkage factor to correct regression coefficients of the original model) was 0.92 IQR 0.91-0.92, the median explained variance was 31% IQR 0.29-0.32 and the AUC 0.78 IQR 0.77-0.78. This means that after corrections for over-optimism, both the performance and discriminative properties of the model remained good. Results are shown in Table 3.

DISCUSSION

This study showed that the Step-Dep intervention was not more effective than usual care in the prevention of MDD at two years of follow-up. The risk of incident MDD during two years of follow-up among patients with DM2 and/or CHD and subthreshold depression, was increased by higher baseline levels of anxiety and depression, the presence of more than three chronic diseases and having suffered a stressful life-event in the past year. This risk was not influenced by a stepped-care intervention aimed at preventing MDD.

Our findings have to be viewed in the context of strengths and limitations of this study. Strengths are its relatively long follow-up with frequent outcome measurements and low drop-out rates. In addition, missing values were imputed using multiple imputation techniques. We only used predictors that are readily available or easily obtained in practice, which enhances the practical use of the model in primary care consultations. Furthermore, testing a multivariable model instead of single factors appointed only the most relevant predictors, which rendered a simple model that is manageable in its use. There were limitations to this study. First, the study population was relatively small, which might have caused over-optimism of the prediction model. This means that it predicts the outcome better in the sample used to develop the model than in new samples, potentially restricting its external validity. However, an internal validation procedure with bootstrapping techniques showed that this risk was minor. Second, we used data derived from a RCT instead of a cohort, which potentially limits the generalizability of our results. Third, we evaluated a limited number of predictors in this study and genetic and other biological risk indicators, for example, were not included. This was due to the relatively small population size and our pre-selection criteria for potential predictors: predictors had to be both identified before in multiple studies and easily obtainable in GP practice . Finally, in this study, the use of the PHQ-9 with a cut-off score of 10 or more rendered a higher cumulative incidence of depression than the MINI. This could be explained by the fact that the PHQ-9 was measured more frequently than the MINI. Also, the PHQ-9 was self-reported instead of administered with a diagnostic interview by a trained research assistant. However, it is possible that depression was sometimes over-diagnosed with the PHQ-9 due to potential overlap of (somatic) symptoms of the chronic disease and those of depression[55].

Our multivariable model consisted of four predictors of MDD incidence. Firstly, baseline depression severity level is the most frequently found and often strongest predictor of incident depression in other studies in patients with DM2[21,22,25] or CHD[28,31,32]. In line with these findings, in our model a clinically relevant baseline difference in depressive symptoms of five points on the PHQ-9, translated to an almost five times increased risk of developing a MDD during two years. This factor was used as a continuous variable in which the severity level predicts the occurrence of a depressive episode, which supports the concept of a gradual risk of depression. Secondly, the anxiety level at baseline was an important predictor of MDD. Anxiety has been frequently appointed as an important risk factor for depression in DM2[21] and CHD populations[30,31]. Predictors are not necessarily etiological factors[56]. Nonetheless, as anxiety is also known for its high comorbidity with depression, the assumption that reducing anxiety will have a positive effect on depressive symptoms and MDD incidence seems defendable. Thirdly, the risk the occurrence of stressful life-events pose, has been demonstrated before in patients with CHD[28]. Although most of our knowledge on the role of stressful life-events as predictors of depression cover a short period of time[57], more recent research has shown their long-term risk[58]. This would imply that healthcare providers should not only be temporarily alert on the negative influence on mental health of stressful life-events, but should also be aware of deferred effects. Fourthly, the presence of more than three chronic diseases was identified as a predictor of MDD in our study, in concordance with results in a DM2 population of Fisher et al.[24] Interestingly, the presence of either DM2, CHD or both was not a predictor in our study, which suggests that these patients are at the same risk of incident depression. As all included patients in Fisher's and our study had at least one chronic disease, a discrimination between the predictive values of no chronic disease versus only one versus multiple chronic diseases could not be made. The specific importance of an increased number of diseases as opposed to the risk of a chronic disease has also been demonstrated previously in a primary care population with subthreshold depression[59] and several elderly populations[60]. Why the number of diseases would matter in itself, can perhaps be understood from findings from qualitative interviews. Step-Dep patients explained that chronic diseases indirectly lead to depression, as they diminish future perspectives and cause disability[40], which might be subjective to a certain "threshold" burden of disease. Finally, in contrast to findings in multiple other studies, female sex[24,27,29,31]and a history of depression[25,27,29] did not predict incident MDD in our study. These factors were also not univariately associated with incident depression in our data. A

history of depression was self-reported in our study. Perhaps patients over-reported this, as it was not required that they received treatment for this depressive episode, which might explain the lack of an univariate correlation with incident depression.

The model rendered in this study had good discriminative properties with an AUC of 0.80 with the use of only four predictors that are relatively easily obtained by the GP. This makes this prediction model practically viable. It could assist as a tool to both improve the (early) recognition of depression in primary care patients with DM2 and/or CHD and indicate which patients need further care. As chronic care in the Netherlands is being delegated more and more to primary care practice nurses, such a tool might prove useful in their and the GPs' regular check-ups. In practice, this would not only entail that in patients with DM2 and/or CHD, GPs and practice nurses standardly inquire about symptoms of depression and anxiety during regular checkups, but also that in those with multiple chronic diseases next to their DM2 or CHD, who suffered a recent stressful life-event, the presence and course of depressive and anxiety symptoms are assessed and monitored over time with, for example, the PHQ-9 and HADS. Whenever depressive or anxiety symptoms are clinically severely elevated or significantly deteriorate over time, treatment should be offered according to the patients' need for care. By reducing both depressive and anxiety symptoms, perhaps MDD and its negative consequences can be averted.

Future research should focus on the external validation to test the generalizability of our results, for example on DM2 and/or CHD patients without subthreshold depression, or outside the Dutch setting. Subsequently, studies are required to investigate the influence of the prediction model on decision making and patient outcomes. Consecutively, future research should evaluate whether the suggested enhanced vigilance strategies in combination with depression prevention programs that only target those with all four indicated predictors present and aim to reduce both anxiety and depressive symptoms, are cost-effective[61].

REFERENCES

1. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 2006;3:2011–30.

2. WHO. World Health Statistics 2017: Monitoring health for the SDGs. 2017.

3. Roy T, Lloyd CE. Epidemiology of depression and diabetes: A systematic review. J. Affect. Disord.; 2012;142:S8–21.

4. Rudisch B, Nemeroff CB. Epidemiology of comorbid coronary artery disease and depression. Biol. Psychiatry. 2003;54:227–40.

5. Lin EHB, Katon W, Von Korff M, Rutter C, Simon GE, Oliver M, et al. Relationship of depression and diabetes self-care, medication adherence, and preventive care. Diabetes Care. 2004;27:2154–60.

6. Gehi A, Haas D, Pipkin S. Depression and medication adherence in outpatients with coronary heart disease. Arch Intern Med. 2005;165:2508–13.

7. Ali S. The association between depression and health-related quality of life in people with type 2 diabetes: a systematic literature review. Diabetes. Metab. Res. Rev. 2010;26:75–89.

8. Rutledge T, Bittner V, Olson MB, Linke SE, Cornell CE, Eteiba W, et al. Depression and Cardiovascular Health Care Costs Among Women With Suspected Myocardial Ischemia (Women's Ischemia Syndrome Evaluation) Study. Jac. American College of Cardiology Foundation; 2009;53:176–83.

9. Bosmans JE, Adriaanse MC. Outpatient costs in pharmaceutically treated diabetes patients with and without a diagnosis of depression in a Dutch primary care setting. BMC Health Serv. Res. 2012;12:46.

10. Katon W, Lin EHB, Von Korff M, Ciechanowski P, Ludman E, Young B, et al. Integrating depression and chronic disease care among patients with diabetes and/or coronary heart disease: The design of the TEAMcare study. Contemp. Clin. Trials.; 2010;31:312–22.

11. Sullivan M, O'Connor P, Feeney P, Hire D, Simmons DL, Raisch D, et al. Depression Predicts All-Cause Mortality. Diabetes Care. 2012;35:1708–15.

12. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. Lancet; 2009;374:609–19.

13. National Collaborating Centre for Mental Health. Depression in adults with a chronic physical health problem. The NICE Guideline of Treatment and Management . 2010.

14. Chisholm D, Sanderson K, Ayuso-Mateos JL, Saxena S. Reducing the global burden of depression: Population-level analysis of intervention cost-effectiveness in 14 world regions. Br. J. Psychiatry. 2004;184:393–403.

15. Cuijpers P, van Straten A, Smit F, Mihalopoulos C, Beekman A. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. Am. J. Psychiatry. 2008;165:1272–80.

16. van Zoonen K, Buntrock C, Ebert DD, Smit F, Reynolds CF, Beekman ATF, et al. Preventing the onset of major depressive disorder: A meta-analytic review of psychological interventions. Int. J. Epidemiol. 2014;43:318–29.

17. Bower P, Gilbody S. Stepped care in psychological therapies: Access, effectiveness and efficiency. Narrative literature review. Br. J. Psychiatry. 2005;186:11–7.

18. van Dijk SEM, Pols AD, Adriaanse MC, Bosmans JE, Elders PJM, van Marwijk HWJ, et al. Cost-effectiveness of a stepped-care intervention to prevent major depression in patients with type 2 diabetes mellitus and/or coronary heart disease and subthreshold depression: design of a cluster-randomized controlled trial. BMC Psychiatry. 2013;13:128.

19. Davidson SK, Harris MG, Dowrick CF, Wachtler CA, Pirkis J, Gunn JM. Mental health interventions and future major depression among primary care patients with subthreshold depression. J. Affect. Disord.; 2015;177:65–73.

20. Pols AD, Van Dijk SE, Bosmans JE, Hoekstra T, Van Marwijk HWJ, Van Tulder MW, et al. Effectiveness of a stepped-care intervention to prevent major depression in patients with type 2 diabetes mellitus and/or coronary heart disease and subthreshold depression: A pragmatic cluster randomized controlled trial. PLoS One. 2017;12.

21. Bot M, Pouwer F, Ormel J, Slaets JPJ, de Jonge P. Predictors of incident major depression in diabetic outpatients with subthreshold depression. Diabet. Med. 2010;27:1295–301.

22. Pibernik-Okanovic M, Begic D, Peros K, Szabo S, Metelko Z. Psychosocial factors contributing to persistent depressive symptoms in type 2 diabetic patients: a Croatian survey from the European Depression in Diabetes Research Consortium. J. Diabetes Complications. 2008;22:246–53.

23. Badawi G, Pagé V, Smith KJ, Gariépy G, Malla A, Wang J, et al. Self-rated health: A predictor for the three year incidence of major depression in individuals with Type II diabetes. J. Affect. Disord. 2013;145:100–5.

24. Fisher L, Skaff MM, Mullan JT, Arean P, Glasgow R, Masharani U. A longitudinal study of affective and anxiety disorders, depressive affect and diabetes distress in adults with type 2 diabetes. Diabet. Med. 2008;25:1096–101.

25. Katon W, Ph JR, M EHBLMD, D SRHM, M PCMD, Evette J Ludman Ph, et al. Depression and Diabetes: Factors Associated With Major Depression at Five-Year Follow-Up. Psychosomatics. 2011;50:570–9.

26. Bot M, Pouwer F, Ormel J, Slaets JPJ, de Jonge P. Predictors of incident major depression in diabetic outpatients with subthreshold depression. Diabet. Med. England; 2010;27:1295–301.

27. Nefs G, Pouwer F, Denollet J, Pop V. The course of depressive symptoms in primary care patients with type 2 diabetes: Results from the Diabetes, Depression, Type D Personality Zuidoost-Brabant (DiaDDZoB) Study. Diabetologia. 2012;55:608–16.

28. Doyle F, McGee H, Delaney M, Motterlini N, Conroy R. Depressive vulnerabilities predict depression status and trajectories of depression over 1 year in persons with acute coronary syndrome. Gen. Hosp. Psychiatry.; 2011;33:224–31.

29. Spijkerman TA, Van Den Brink RHS, Jansen JHC, Crijns HJGM, Ormel J. Who is at risk of post-MI depressive symptoms? J. Psychosom. Res. 2005;58:425–32.

30. Pedersen SS, Denollet J, van Gestel YRBM, Serruys PW, van Domburg RT. Clustering of psychosocial risk factors enhances the risk of depressive symptoms 12-months post percutaneous coronary intervention. Eur. J. Cardiovasc.

Prev. Rehabil. 2008;15:203–9.

31. Ossola P, Paglia F, Pelosi A, De Panfilis C, Conte G, Tonna M, et al. Risk factors for incident depression in patients at first acute coronary syndrome. Psychiatry Res. 2015;228:448–53.

32. Kang H, Stewart R, Bae K, Kim S, Shin I, Hong Y, et al. Predictors of depressive disorder following acute coronary syndrome: Results from K-DEPACS and EsDEPACS. J. Affect. Disord.; 2015;181:1–8.

33. Kroenke K SR. The PHQ-9: a new depression diagnostic and severity measure. Psychiatr Ann. 2002;32:509–515.

34. Lamers F, Jonkers CCM, Bosma H, Penninx BWJH, Knottnerus JA, van Eijk JTM. Summed score of the Patient Health Questionnaire-9 was a reliable and valid method for depression screening in chronically ill elderly patients. J. Clin. Epidemiol. 2008;61:679–87.

35. Sheehan DV, Lecrubier Y SK. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatr. 1998;59:22–33.

36. Van Vliet IM, De Beurs E. Het Mini Internationaal Neuropsychiatrisch Interview (MINI): Een kort gestructureerd diagnostisch psychiatrisch interview voor DSM-IV-en ICD-10-stoornissen. Tijdschr. Psychiatr. 2007;49:393–7.

37. Meader N, Mitchell AJ, Chew-graham C, Goldberg D, Rizzo M, Bird V, et al. Case identification of depression in patients with chronic physical health problems: a diagnostic accuracy meta-analysis of 113 studies. 2011;808–20.

38. Moriarty AS, Gilbody S, McMillan D, Manea L. Screening and case finding for major depressive disorder using the Patient Health Questionnaire (PHQ-9): A meta-analysis. Gen. Hosp. Psychiatry; 2015;37:567–76.

39. van der Zwaan GL, van Dijk SEM, Adriaanse MC, van Marwijk HWJ, van Tulder MW, Pols AD, et al. Diagnostic accuracy of the Patient Health Questionnaire-9 for assessment of depression in type II diabetes mellitus and/or coronary heart disease in primary care. J. Affect. Disord. 2015;190:68–74.

40. Pols A, Schipper K, Overkamp D, van Marwijk H, van Tulder M, Adriaanse M. Patients' and practice nurses' perceptions of depression in patients with diabetes type 2 and/or coronary heart disease screened for

subthreshold depression. submitted.

41. Spinhoven P, Ormel J, Sloekers PP, Kempen GI, Speckens AE, Van Hemert AM. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. Psychol Med. 1997;27:363–70.

42. Robins LN, Helzer JE, Croughan J, Ratcliff KS. National Institute of Mental Health Diagnostic Interview Schedule. Arch Gen Psychiatry. 1981;38:381–9.

43. Kriegsman DMW, Penninx BWJH, Van Eijk JTM, Boeke a. JP, Deeg DJH. Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly. A study on the accuracy of patients' self-reports and on determinants of inaccuracy. J. Clin. Epidemiol. 1996;49:1407–17.

44. Twisk J. Different Methods to Analyse the Results of a Randomized Controlled Trial with More Than One Follow-Up Measurement. In: K. van Montfoort, J. Oud & WG, editor. Dev. Stat. Eval. Clin. Trials.; 2014. p. 177–93.

45. McCulloch CE NJ. Generalized Linear Mixed Models. Encyclopedia of Biostatistics. John Wiley & Sons; 2005.

46. Twisk JW. Applied longitudinal data analysis for epidemiology: a practical guide. Cambridge University Press; 2013.

47. Ali S, Stone MA, Peters JL, Davies MJ, Khunti K. The prevalence of co-morbid depression in adults with Type 2 diabetes: a systematic review and meta-analysis. Diabet. Med. 2006;23:1165–73.

48. Comijs HC, Nieuwesteeg J, Kok R, van Marwijk HW, van der Mast RC, Naarding P, et al. The two-year course of late-life depression; results from the Netherlands study of depression in older persons. BMC Psychiatry.; 2015;15:1.

49. de Boer MR, Waterlander WE, Kuijper L, Steenhuis I, Twisk J. Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate. Int. J. Behav. Nutr. Phys. Act. 2015;12:4.

50. Van Buuren, S; Groothuis-oudshoorn K. mice : Multivariate Imputation by Chained. J. Stat. Softw. 2011;45.

51. Rubin DB. Multiple imputation for nonresponse in surveys. John Wiley & Sons; 2004.

52. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. Eur. Urol. 2015;67:1142–51.

53. Heymans MW, Buuren S Van, Knol DL, Van W, Vet HCW De. Variable selection under multiple imputation using the bootstrap in a prognostic study. 2007;10:1–10.

54. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. BMC Med. Res. Methodol. 2009;9:57.

55. Reddy P, Philpot B, Ford D, Dunbar JA. Identification of depression in diabetes: the efficacy of PHQ-9 and HADS-D. Br. J. Gen. Pract. England; 2010;60:e239-45.

56. Moons K, Royston P, Vergouwe Y, Grobbee D, Altman D. Prognosis and prognostic research: what, why, and how? BMJ. 2009;338.

57. KENDLER, KENNETH S.; KARKOWSKI, LAURA M.; PRESCOTT CA. Stressful Life Events and Major Depression: Risk Period, Long-Term Contextual Threat, and Diagnostic Specificity. J. Nerv. Ment. Dis. 1998;186:661–9.

58. Assari S, Lankarani MM. Stressful life events and risk of depression 25 years later: Race and gender differences. Front. Public Heal. 2016;4:49.

59. Cuijpers P, Smit F, Willemse G. Predicting the onset of major depression in subjects with subthreshold depression in primary care: A prospective study. Acta Psychiatr. Scand. 2005;111:133–8.

60. Vink D, Aartsen MJ, Schoevers RA. Risk factors for anxiety and depression in the elderly: A review. J. Affect. Disord. 2008;106:29–44.

61. Steyerberg E, Moons KGM, van der Windt D, Hayden J, Perel P, Schroter S, et al. Prognosis research strategy PROGRESS series 3: prognostic model research. PLoS Med. 2013;10:e1001381.

Table 1 Patients' baseline characteristics at baseline in intervention group, care as usual group and total sample

Characteristics	Total sample (N=236)	Intervention (N=96)	Care as usual (N=140)
Female	107/236 (45.3)	42/96 (43.8)	65/140 (46.4)
Age, mean (SD)	67.5 (10.0)	67.8 (9.2)	67.3 (10.5)
Stressful life-event	112/210 (53.3)	48/89 (53.9)	64/121 (52.9)
Positive history of depression	113/210 (53.8)	54/89 (60.7)	59/121 (48.8)
ICPC diagnosis DM2 and/or CHD			
Diabetes Mellitus type 2 (DM2)	88/236 (37.3)	38/96 (39.6)	50/140 (35.7)
Coronary Heart Disease (CHD)	86/236 (36.4)	36/96 (37.5)	50/140 (35.7)
DM2 and CHD	62/236 (26.3)	22/96 (22.9)	40/140 (28.6)
More than 3 chronic diseases	98/210 (46.7)	38/89 (42.7)	60/121 (49.6)
PHQ-9 at baseline, mean (SD)	9.4 (3.2)	9.5 (3.1)	9.3 (3.2)
Anxiety HADS, mean (SD)	6.5 (3.8)	6.9 (3.7)	6.3 (3.9)
Depression HADS, mean (SD)	6.5 (3.8)	6.9 (3.9)	6.1 (3.7)
Marital status			
Married/living together	122/220 (55.5)	55 (61.1)	67/130 (51.5)
Single/divorced/widowed	98/220 (44.5)	35 (38.9)	63/130 (48.5)
Both parents born in the Netherlands	186/220 (84.5)	74/90 (82.2)	112/130 (86.2)
Rural residential area	99/236 (41.9)	42 (43.8)	57/140 (40.7)
Unemployed/sick	26/220 (11.8)	12/90 (13.3)	14/130 (10.8)
Level of education			
Low	89/220 (40.5)	33/90 (36.7)	56/130 (43.1)
Average	60/220 (27.3)	22/90 (24.4)	38/130 (29.2)
High	71/220 (32.3)	35/90 (38.9)	36/130 (27.7)
Current smoker	39/219 (17.8)	16/90 (17.8)	23/129 (17.8)
Alcohol use above norm	63/219 (28.8)	29/90 (32.2)	34/129(26.4)
Exercise under norm	141/219 (64.4)	56/90 (62.2)	85/129 (65.9)
BMI, mean (SD)	28.9 (6.1)	29.4 (6.8)	28.5 (5.6)
Locus of Control, mean (SD)	7.9 (4.2)	8.3 (4.2)	7.6 (4.1)
Social support, mean (SD)	36.3 (9.2)	35.8 (9.0)	36.7 (9.5)
Dysthymia	13/236 (5.5)	6/96 (6.3)	7/140 (5.0)
Onset of depression after age of 55	101/210 (48.1)	38/89 (42.7)	63/121 (52.1)

Numbers are percentages unless stated otherwise; Abbreviations: BMI = Body Mass Index; PHQ-9, Patient Health Questionnaire-9; HADS, Hospital Anxiety and Depression Scale; SD, Standard Deviation.

Table 2 Results of the mixed model and GEE long-term effectiveness analyses

Cumulative incidence of depression (n/N) %	Intervention	Care as usual	Corrected analyses*		Crude analyses	
Baseline	0	0	**OR (95%CI)**	**P-value**	**OR (95%CI)**	**P-value**
T6	(5/84) 6.0	(10/125) 8.0	0.82 (0.19; 3.51)	0.79	0.90 (0.32; 2.50)	0.84
T12	(9/82) 11.0	(12/118) 10.2	1.44 (0.46; 4.47)	0.53	1.20 (0.49; 2.92)	0.70
T24	(13/77) 16.9	(17/105) 16.2	1.23 (0.50; 3.02)	0.66	1.11 (0.51; 2.44)	0.79
Overall effect	n.a	n.a	1.37 (0.52;3.55)	0.52	1.11 (0.49;2.49)	0.80
PHQ mean (SD)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	9.53 (3.14)	9.28 (3.23)	**B (95%CI)**	**P-value**	**B (95%CI)**	**P-value**
T3	6.68 (4.55)	6.58 (4.21)	-0.39 (-1.52; 0.74)	0.50	-0.03 (-1.17; 1.11)	0.96
T6	6.10 (4.43)	6.12 (4.41)	-0.37 (-1.50; 0.76)	0.52	-0.17 (-1.30; 0.95)	0.76
T9	6.28 (4.31)	6.46 (4.51)	-0.48 (-1.62; 0.65)	0.40	-0.40 (-1.53; 0.73)	0.49
T12	6.60 (5.23)	6.29 (4.46)	-0.09 (-1.20; 1.02)	0.88	-0.03 (-1.13; 1.07)	0.96
T24	5.81 (4.76)	5.15 (4.33)	0.00 (-1.18; 1.19)	0.88	0.02 (-1.15; 1.19)	0.97
Overall effect	n.a	n.a	0.29 (-1.15; 0.58)	0.52	-0.13 (-0.99; 0.73)	0.77
Perceived recovery (%)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	n.a	n.a	**OR (95%CI)**	**P-value**	**OR (95%CI)**	**P-value**
T3	40.3%	49.5%	0.78 (0.42; 1.45)	0.44	0.64 (0.36; 1.15)	0.14
T6	48.8%	45.5%	1.46 (0.79; 2.69)	0.23	1.15 (0.65; 2.02)	0.64
T9	55.0%	48.7%	1.47 (0.79; 2.75)	0.22	1.30 (0.74; 2.30)	0.91
T12	55.6%	58.1%	1.04 (0.56; 1.92)	0.91	0.91 (0.51; 1.61)	0.74
T24	68.0%	57.1%	2.38 (1.21; 4.67)	0.01	2.04 (1.08; 3.87)	0.03
Overall effect	n.a	n.a	1.32 (0.87; 2.00)	0.19	1.10 (0.75; 1.62)	0.61
HADS-A mean (SD)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	6.91 (3.74)	6.25 (3.90)	**B (95%CI)**	**P-value**	**B (95%CI)**	**P-value**
T3	6.35 (4.04)	6.29 (3.97)	-0.27 (-1.13; 0.60)	0.54	-0.13 (-1.00; 0.74)	0.76
T6	5.70 (4.10)	6.63 (4.00)	-1.04 (-1.91; -0.18)	0.02	-1.04 (-1.91; -0.18)	0.02
T9	6.16 (4.24)	6.03 (4.04)	-0.49 (-1.35; 0.38)	0.27	-0.45 (-1.31; 0.42)	0.31
T12	5.77 (4.69)	5.83 (3.99)	-0.50 (-1.37; 0.38)	0.27	-0.43 (-1.31; 0.44)	0.33
T24	5.45 (4.46)	5.06 (3.90)	-0.59 (-1.50; 0.31)	0.20	-0.48 (-1.38; 0.43)	0.30
Overall effect	n.a	n.a	-0.59 (-1.23; 0.06)	0.08	-0.52 (-1.17; 0.13)	0.12
HADS-D mean (SD)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	6.93 (3.87)	6.11 (3.73)	**B (95%CI)**	**P-value**	**B (95%CI)**	**P-value**
T3	6.14 (4.16)	6.21 (3.87)	-0.26 (-1.12; 0.60)	0.55	-0.29 (-1.15; 0.56)	0.51
T6	5.82 (3.79)	5.75 (4.03)	-0.22 (-1.07; 0.64)	0.62	-0.32 (-1.18; 0.53)	0.46
T9	6.36 (4.04)	6.07 (4.08)	-0.21 (-1.06; 0.65)	0.63	-0.24 (-1.09; 0.61)	0.58
T12	6.09 (4.20)	6.11 (4.22)	-0.41 (-1.27; 0.46)	0.36	-0.50 (-1.36; 0.36)	0.26
T24	5.59 (4.66)	4.92 (3.90)	-0.41 (-1.30; 0.48)	0.37	-0.48 (-1.37; 0.41)	0.29
Overall effect	n.a	n.a	-0.30 (-0.94; 0.33)	0.35	-0.37 (-1.00; 0.26)	0.25

Abbreviations: 95%CI, 95% Confidence Interval; HADS-A, Hospital Anxiety and Depression Scale-Anxiety; HADS-D, Hospital Anxiety and Depression Scale-Depression; n.a, not applicable; PHQ-9, Patient Health Questionnaire-9;

*Corrected for: baseline values of the outcome, age, gender, marital status, employment status, level of education, co-existence of DM2 and CHD, alcohol use, number of depressive episodes in history and age of onset of depression. The baseline value of the outcome is not added as an extra variable in the corrected analyses of the overall effects since it is already incorporated in the crude overall analyses.

Table 3 Multivariable prediction model of incident depression during two-year follow-up

Predictor	RC	OR	95% CI	P-value
Female sex		-	-	-
Age		-	-	-
Somatic disorder DM2 CHD DM2and CHD		-	-	-
History of depression		-	-	-
Baseline depression scores	0.32 p.p.i.	1.37	1.20; 1.55	0.00
Baseline anxiety scores	0.12 p.p.i.	1.13	1.02; 1.25	0.01
Stressful life-event in past year	0.74	2.10	1.02; 4.32	0.04
>3 chronic illnesses	0.78	2.19	1.12; 4.25	0.02
Randomization status I vs C	0.14	1.15	0.58; 2.29	0.68

RC regression coefficient; p.p.i. per point increase; 95% CI 95% confidence interval; OR odds ratio, an OR > 1 reflects a higher probability the outcome an incident depression and an OR < 1 reflects a lower probability compared with the reference category. OR estimated after multiple imputation (n = 25 datasets) with p-value of 0.157. Linear predictor corrected after bootstrapping = -4.1147 + 0.131* Randomization status + 0.7167* >3 chronic illnesses + 0.680* stressful life-event in past year + 0.1118* baseline anxiety scores + 0.2868* baseline depression scores

SUPPLEMENTARY INFORMATION

Contributors

AP constructed the design of this study, performed all statistical analyses and drafted the manuscript. MCA, MvT, HvM constructed the design of the study and revised the manuscript. JB and SvD constructed the design of the Step-Dep study and revised the manuscript. MH collaborated on the statistical analyses and revised the manuscript. The final manuscript was read and approved by all authors.

Funding

This study is funded by ZonMw, the Netherlands Organisation for Health Research and Development (project number 80-82310-97-12110). The sponsor had no role in the design and conduct of the present study or in the writing of the manuscript.

Competing Interests

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Trial registration and ethical approval

The study was performed in accordance with the declaration of Helsinki (2008) and the Dutch Medical Research involving Human Subjects Act (WMO). The protocol was approved by the medical ethics committee of the VU University Medical Centre (NL39261.029.12, registration number 2012/223), and registered in the Dutch Trial Register (registration number 3715).

Data sharing

Full dataset and statistical code is available from the corresponding author. Consent was not obtained but the presented data are anonymised and risk of identification is low.

Transparency

The corresponding author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; no important aspects of this study have been omitted; and any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Acknowledgements

The authors would like to thank Marcella van der Linden, Lucca Vledder and Mieke Schlattmann for their contribution in the data collection for this study and Jos Twisk for his help in the long-term effectiveness analyses. We also would like to thank all the participating general practices and the research networks of general practitioners (ANH, THOON and LEON) for their participation and collaboration in the implementation and execution of the study. Furthermore, this study has been possible thanks to all Step-Dep participants.

SUPPLEMENTARY FILES

S1 Original protocol

S2 TRIPOD statement checklist

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	4-5
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	5
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	6
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	6
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	6
	5b	D;V	Describe eligibility criteria for participants.	6
	5c	D;V	Give details of treatments received, if relevant.	7
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	6
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	6
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	7
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	6
Sample size	8	D;V	Explain how the study size was arrived at.	Ref protocol
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	8
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	7-8
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	8-9
	10c	V	For validation, describe how the predictions were calculated.	n.a.
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	9
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	n.a.
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	n.a.
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	n.a.
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	10
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	10, table 1
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	n.a.
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	n.a.
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	n.a.
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	Table 3
	15b	D	Explain how to the use the prediction model.	10-11
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	10-11
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	n.a.
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	12
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	n.a.
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	12-13
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	13-14
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	24-26
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	24

TRIPOD Checklist: Prediction Model Development and Validation

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

BMJ Open

Two-year effectiveness of a stepped-care depression prevention intervention and predictors of incident depression in primary care patients with diabetes type 2 and/or coronary heart disease and subthreshold depression; data from the Step-Dep cluster randomized controlled trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020412.R1
Article Type:	Research
Date Submitted by the Author:	26-Feb-2018
Complete List of Authors:	Pols, Alide; VU University, Department of Health Sciences Adriaanse, Marcel; Institute of Health Sciences, Vrije Universiteit Amsterdam Tulder, Maurits; University of Amsterdam, Health Sciences Heymans, Martijn; VU University Medical Center, Department of Epidemiology and Biostatistics; VU University, Department of Health Sciences Section Methodology and Applied Biostatistics Faculty of Earth and Life Sciences Bosmans, J; VU University Amsterdam, Department of Health Sciences van Dijk, Susan; Vrije Universiteit Amsterdam, Department of Health Sciences Van Marwijk, H; University of Brighton, Division of Primary Care and Public Health, Brighton and Sussex Medical School, Mayfield House
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Cardiovascular medicine, Diabetes and endocrinology, Mental health, General practice / Family practice
Keywords:	Depression & mood disorders < PSYCHIATRY, DIABETES & ENDOCRINOLOGY, Coronary heart disease < CARDIOLOGY, PREVENTIVE MEDICINE

Two-year effectiveness of a stepped-care depression prevention intervention and predictors of incident depression in primary care patients with diabetes type 2 and/or coronary heart disease and subthreshold depression; data from the Step-Dep cluster randomized controlled trial

Alide D Pols*[1,2], Marcel C Adriaanse[1], Maurits W van Tulder[1], Martijn W Heymans[3], Judith E Bosmans[1], Susan E van Dijk[1], Harm W J van Marwijk[2,4]

[1] Department of Health Sciences, Amsterdam Public Health research institute, Faculty of Science, Vrije Universiteit, Amsterdam, The Netherlands

[2] Department of General Practice & Elderly Care Medicine, Amsterdam Public Health research institute, VU University Medical Centre, Amsterdam, The Netherlands

[3] Department of Clinical Epidemiology and Biostatistics VU Medical Centre Amsterdam, The Netherlands

[4] Division of Primary Care and Public Health, Brighton and Sussex Medical School, Mayfield House, University of Brighton, Brighton, United Kingdom

* Corresponding author: a.d.pols@vu.nl

Keywords: major depressive disorder, subthreshold depression, diabetes mellitus type 2, coronary heart disease, effectiveness, stepped care, prediction model

ABSTRACT

Introduction Major depressive disorders (MDD), diabetes mellitus type 2 (DM2) and coronary heart disease (CHD) are leading contributors to the global burden of disease and often co-occur.

Objectives To evaluate the two-year effectiveness of a stepped-care intervention to prevent MDD compared to usual care and to develop a prediction model for incident depression in DM2 and/or CHD patients with subthreshold depression.

Methods Data of 236 Dutch primary care DM2/CHD patients with subthreshold depression (Patient Health Questionnaire 9 (PHQ-9) score ≥6, no current MDD according to the Mini International Neuropsychiatric Interview (DSM-IV criteria)), who participated in the Step-Dep trial were used. A PHQ-9 score of ≥10 at minimally one measurement during follow-up (at 3, 6, 9, 12 and 24 months) was used to determine the cumulative incidence of MDD. Potential demographic and psychological predictors were measured at baseline via web-based self-reported questionnaires and evaluated using a multivariable logistic regression model. Model performance was assessed with the Hosmer–Lemeshow test, Nagelkerke's R^2 explained variance and Area Under the Receiver Operating Characteristic curve (AUC). Bootstrapping techniques were used to internally validate our model.

Results 192 patients (81%) were available at two-year follow-up. The cumulative incidence of MDD was 97/192 (51%). There was no statistically significant overall treatment effect over 24 months of the intervention (OR 1.37; 95% CI 0.52; 3.55). Baseline levels of anxiety, depression, the presence of >3 chronic diseases and stressful life-events predicted the incidence of MDD (AUC 0.80 interquartile range (IQR) 0.79-0.80; Nagelkerke's R^2 0.34 IQR 0.33-0.36).

Conclusion A model with four factors predicted depression incidence during two-year follow-up in patients with DM2/CHD accurately, based on the AUC. The Step-Dep intervention did not influence the incidence of MDD. Future depression prevention programs should target patients with these four predictors present, and aim to reduce both anxiety and depressive symptoms.

TRIAL REGISTRATION NUMBER

Dutch Trial Register NTR3715 http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=3715

STRENGTHS AND LIMITATIONS OF THIS STUDY

- This study provides a prediction model of incident MDD in DM2 and/or CHD patients with subthreshold depression, which could assist healthcare providers in its detection and facilitate targeting indicated prevention to highest risk patients
- Only predictors that are readily available or easily obtained in practice were used in the multivariable model, which enhances the practical use of the model
- This study had a relatively long follow-up and outcomes were frequently measured, whereas drop-out rates were relatively low and missing values imputed
- The relatively small study population might have caused over-optimism of the prediction model, but an internal validation procedure with bootstrapping techniques showed that this risk was minor
- Data were derived from a RCT, but statistically non-significant intervention effects for incident MDD at both 12- and 24-months follow-up justify using the Step-Dep population as a cohort

INTRODUCTION

Depression is a major and increasing contributor to the global burden of disease[1], whereas coronary heart disease (CHD) and diabetes mellitus type 2 (DM2) rank among the leading causes of morbidity and mortality worldwide[2]. Comorbid depression in patients with DM2 and/or CHD is common[3,4] and has detrimental effects on self-care and medication adherence[5,6], quality of life[7], health status and increases healthcare costs[8,9] and mortality[10,11]. Despite its negative impact, many cases of depression go unrecognized in primary care[12], especially in patients with chronic diseases like DM2 and/or CHD[13]. Additionally, about one-third of those recognized and treated does not respond to current approaches, and over half of those who experience a first onset of a major depressive episode will experience one or more recurrences[14].

Given the significant burden of disease of depression, its poor recognition and the limited effect of current treatment options for it, it would be of great value if incident cases could be averted by early detection and preventive treatment of patients at risk ('indicated prevention'). Meta-analyses have shown that preventive psychological interventions can overall reduce the incidence of MDD in comparison to control groups[15,16]. Offering preventive psychological interventions in a stepped-care format could be an efficient approach, as patients start with minimally intensive evidence-based treatments and only those who do not improve adequately, step up to a treatment of higher intensity[17]. Recently, we conducted a randomized controlled trial in which we evaluated whether a pragmatic nurse-led stepped-care program was effective in reducing the incidence of MDD at 12-months of follow-up in comparison with usual care among patients with DM2 and/or CHD and subthreshold depression (Step-Dep)[18]. Subthreshold depression entails clinically relevant depressive symptoms without fulfilling the criteria for MDD and is a known important risk factor for depression[15,19]. We demonstrated that the Step-Dep intervention was not superior to usual care and the overall cumulative incidence of MDD was lower than expected after one year [20]. However, it may be possible that the follow-up period was too restricted to demonstrate the potential health benefits of the stepped-care program over usual care, or the presence of subthreshold depression alone posed a lower than expected prior risk of MDD in our DM2 and/or CHD population.

Identifying additional major risk factors of incident depression in patients with DM2 and/or CHD might facilitate targeting indicated prevention to patients with highest risk, but also potentially aid in its detection. In patients with DM2, several longitudinal studies have been conducted to determine risk factors for comorbid incident depression. However, these studies have rendered heterogeneous results, due to small patient samples (<80 at follow-up)[21,22], analyses of single factors only[23,24], the use of mixed samples of type 1 diabetes and DM2[25], patients with either no MDD at baseline[23,26] or both with and without depression at baseline[22,24,25,27], and differences across community[23,24], primary care[25,27] and secondary care settings[22,26]. In patients with CHD, the only available longitudinal data are derived from studies in patients with acute coronary syndrome followed-up after hospital discharge[28–32]. Predictors that were repeatedly identified in DM2 or CHD studies were: depression severity at baseline[21,22,25,28,31,32], history of depression[25,27,29], female sex[24,27,29,31] and baseline anxiety levels[21,30,31]. However, data of patients with both DM2 and CHD, non-acute CHD or within primary care settings are scarce. The goal of the present study was twofold: (1) to evaluate the two-year effectiveness of a nurse-led stepped-care intervention to prevent MDD as compared to usual care (Step-Dep); and to (2) develop a prediction model for incident depression during two-year follow-up in primary care patients with DM2 and/or CHD and subthreshold depression.

METHODS

Design

Data of the Step-Dep cluster randomized controlled trial were used. Step-Dep was conducted in 27 general practitioner (GP) practices in three regions in the Netherlands (Amsterdam, Leiden, Twente), between January 2013 and November 2016, including recruitment and two years of follow-up. A statistician blinded to the characteristics of the GP practices performed the randomization of GP practices using a computer generated list of random numbers. Randomization was done at the level of the GP practice, which corresponds to the participating practice nurse, to avoid contamination between the treatment groups, and was stratified for size (less or more than 5000 patients). The study was performed in accordance with the declaration of Helsinki (2008) and the Dutch Medical Research involving Human Subjects Act (WMO). The protocol was approved by the medical ethics committee of the VU University Medical Centre (NL39261.029.12, registration number 2012/223), and registered in the Dutch Trial Register (NTR3715 http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=3715). Further details on the methods and design of the Step-Dep study have been published elsewhere[18].

Patient and Public Involvement

Patients were not involved in determining the design, the recruitment to or conduct of the study. The medical ethics committee of the VU University Medical Centre assessed the burden of the intervention and participation in the study in general as acceptable for patients. The burden of and satisfaction with the intervention were assessed in a process evaluation with 15 patients. All patients are thanked in the acknowledgements section. Results of the study will be disseminated by letter to all participants.

Patients

Included patients were aged 18 years or more who had an International Classification of Primary Care (ICPC) diagnosis of DM2 and/or CHD and had subthreshold depression identified by screening. Patients with a Patient Health Questionnaire 9 (PHQ-9; range 0-27 with higher scores indicating more severe depressive symptoms) score of six or higher[33,34], and no major depressive disorder according to the Mini International Neuropsychiatric

Interview (MINI)[35,36], were considered to have subthreshold depression. Exclusion criteria were cognitive impairment, psychotic illnesses, a terminal illness, the use of anti-depressant medication, a history of suicide attempt(s), loss of significant other in the past six months, visual impairment, current pregnancy, bipolar disorder, borderline personality disorder or any difficulties completing written questionnaires or visiting the primary care center. A total of 236 patients gave informed consent to participate.

Outcome measure

The outcome measure used was an incident depression (yes/no) defined as a PHQ-9 score of ≥10 at minimally one moment during follow-up (measured at 3, 6, 9, 12 and 24 months after baseline). The PHQ-9 is a widely used and validated instrument that performs well in patients with chronic medical illnesses both as dichotomous diagnosis of major and minor depression and a continuous severity score[34,37]. A cut-off of ≥10 has been shown to be the optimum cutoff for major depression[38], also in this patient group [39]. PHQ-9 was self-reported with web-based or written questionnaires. When these web-based or written questionnaires were not completed, the PHQ-9 was administered by telephone by trained research-assistants, blinded to randomization status.

Potential predictors

The selection of the potential predictors was based on a thorough literature search. Predictors of incident depression that were identified in multiple studies in patients with DM2 or CHD and are routinely available or easily obtained in daily GP practice were used. Additionally, we chose the presence of multiple chronic diseases[24] and stressful life-events[28] although they were identified in single studies only, as these were also indicated as causes of depression by patients and practice nurses in semi-structured interviews as part of the process evaluation of Step-Dep[40], and age[23].

Apart from GP information system derived data on *sex, age* and *ICPC diagnosis of DM2 and/or CHD*, demographics and psychological factors were measured at baseline via web based (or written if preferred) self-reported questionnaires. To take possible effects of the intervention into account, we included randomization status in the selection models as well. Patients in the intervention arm were offered a stepped care prevention program, and

patients in the control arm received care as usual during one year. The stepped care intervention consisted of four sequential but flexible treatment steps, each lasting three months; 1) watchful waiting, 2) guided self-help, 3) problem solving treatment and 4) referral to a general practitioner. After each step, patients with a persisting PHQ-9 score of six or more were offered the next treatment step of the intervention. *Baseline depression levels* were measured with the PHQ-9[33,34]. *Baseline anxiety levels* were measured with the Hospital Anxiety and Depression Scale Anxiety (HADS-A; range 0-21 with higher scores indicating more severe anxiety)[41]. *History of depression* and s*tressful life-events were* self-reported using a subset of the Diagnostic Interview Schedule (DIS)[42]. *Number of co-morbid chronic illnesses* was measured using the self-reported Dutch Questionnaire Chronic Illnesses[43]. This was dichotomized using the median in our sample: three or less versus more than three chronic diseases.

Statistical analyses

The two-year effectiveness of the intervention on the primary and secondary outcomes was analyzed according to the intention to treat principle. Generalised Estimating Equations (GEE) were used for binary outcome variables, and linear mixed models for longitudinal data were used for continuous outcome variables[44]. For each outcome an overall effect over time and separate effects at different time points were estimated by taking time into account as a categorical variable (with five categories: 0-3 months, 3-6 months, 6-9 months, 9-12 months and 12-24 months of follow-up)[45,46]. The main analyses consisted of fully corrected models that were corrected for baseline values of the respective outcome and additionally included the covariates gender[47], age[48], and any other possible confounding variable on which the treatment groups differed at baseline (marital status, employment status, level of education, co-existence of DM2 and CHD, alcohol use, number of depressive episodes in history and age of onset of depression), based on absolute baseline differences judged by the researchers rather than statistical testing[49]. For these analyses, STATA version 14 was used.

Missing data were imputed using multiple imputation according to the Multivariate Imputation by Chained Equations (MICE) algorithm[50] in SPSS version 23. For the imputations, missing at random (MAR) was assumed. Variables that were associated with missing data and variables that were associated with the outcome, were identified and included in the imputation model. Also, all variables in the analysis model (potential predictors and

outcome) were included. The number of imputed datasets was 25 based on the proportion of cases with incomplete measurements; 24%. The subsequent analyses were performed on pooled data according to Rubin's rules[51].

Prediction model

We created a multivariable logistic regression model in SPSS 23 from the baseline variables estimating the probability of having at least one major depression (PHQ ≥10) during the two-year assessment. To calculate the number of potential predictors for developing the prediction model, we used the criterion of 10 events per variable. Continuous variables were checked for linearity with the outcome using spline regression curves and linearity was confirmed. All variables were entered into the logistic model and tested for statistical significance in the presence of the total set of predictors. Individually, the least significant predictor (P-value>0.157, as recommended in the TRIPOD statement, [52], Wald statistic) was removed, and the model was refit (backward selection). Randomization status was maintained in the model. This was repeated until we reached a statistical model that only included statistically significant predictors. This was repeated with p-values of 0.05. We also compared the results with complete case analysis (CCA), i.e., all patients with missing data were excluded from the analyses.

We checked the performance of the model with regard to the goodness of fit (Hosmer–Lemeshow test), the explained variation and the discriminative ability of the model. The Nagelkerke's R^2 explained variation is the extent to which the outcome can be predicted by the predictors in the model in current datasets. The discriminative ability is reflected by the area under the receiver operating characteristics curve (AUC). Bootstrapping techniques were used to internally validate our model, i.e., to simulate the performance with respect to the explained variance and the AUC in comparable patient datasets[53]. After that, we calculated the linear predictor of the bootstrapped model with an adjusted intercept and regression coefficients corrected for the shrinkage factor. Performance measures were assessed in each imputed dataset and results were summarized using median values [54]. All analyses were done with SPSS version 23.0 and R software.

RESULTS

Participants

The baseline characteristics of the study population are presented in Table 1. Of the 236 patients included in Step-Dep, 192 patients (81%) completed two years of follow-up. A flowchart of participants through the first 12 months of the Step-Dep study has been published elsewhere[20]. At 24 months of follow up 18 additional patients dropped out (two for unknown motives, seven due to time considerations, four were deceased, three too frail, two unable to contact). We compared the baseline characteristics of patients with missing data to those without. Patients with missing data were more often living alone (61% vs 41%), but no other differences between these groups were found.

There was no statistically significant overall treatment effect over 24 months of the intervention (OR 1.37; 95% CI 0.52 to 3.55), nor at any of the time-points. There were no significant differences in PHQ-9 scores between the study groups at any time-point and the course of PHQ-9 scores over time did not differ significantly between the groups. Results are shown in Table 2. The statistically non-significant intervention effects for incident MDD at both 12-months[20] and 24-months of follow-up justify using the Step-Dep population as a cohort.

Prediction model

The cumulative incidence during two-year follow-up was 97/192 (51%). The multivariable models using p=0.05 and p=0.157[52] were identical. The final model consisted of four predictors: level of anxiety, level of depression, presence of more than three chronic diseases and having suffered a stressful life-event in the past year. This model performed well (Hosmer–Lemeshow test p=0.12 and median of pooled Nagelkerke's R^2 explained variance 0.34 interquartile range (IQR) 0.33-0.36) with good discriminative properties (median of the pooled AUC 0.80 IQR 0.79-0.80). In a CCA with p=0.05, the same predictors remained. In a CCA using p=0.157 [52], the categorical variable DM2/CHD/both also remained.

The risk of an incident MDD during two years of follow-up more than doubled when either more than three chronic diseases were present or a patient had suffered a stressful life-event in the past year. Both higher

depression and anxiety levels at baseline increased the risk of MDD with each incremental point on the PHQ-9 of HADS scales respectively. One point higher on the PHQ-9 at baseline, resulted in a 1.37 higher risk of developing MDD during two years, compared to 1.13 for increasing anxiety levels. With regard to the internal validation of the model, the calibration slope (or shrinkage factor to correct regression coefficients of the original model) was 0.92 IQR 0.91-0.92, the median explained variance was 31% IQR 0.29-0.32 and the AUC 0.78 IQR 0.77-0.78. This means that after corrections for over-optimism, both the performance and discriminative properties of the model remained good. Results are shown in Table 3.

DISCUSSION

This study showed that the Step-Dep intervention was not more effective than usual care in the prevention of MDD at two years of follow-up. The risk of incident MDD during two years of follow-up among patients with DM2 and/or CHD and subthreshold depression, was increased by higher baseline levels of anxiety and depression, the presence of more than three chronic diseases and having suffered a stressful life-event in the past year. This risk was not influenced by a stepped-care intervention aimed at preventing MDD.

Our findings have to be viewed in the context of strengths and limitations of this study. Strengths are its relatively long follow-up with frequent outcome measurements and low drop-out rates. In addition, missing values were imputed using multiple imputation techniques. We only used predictors that are readily available or easily obtained in practice, which enhances the practical use of the model in primary care consultations. Furthermore, testing a multivariable model instead of single factors appointed only the most relevant predictors, which rendered a simple model that is manageable in its use. There were limitations to this study. First, the study population was relatively small, which might have caused over-optimism of the prediction model. This means that it predicts the outcome better in the sample used to develop the model than in new samples, potentially restricting its external validity. However, an internal validation procedure with bootstrapping techniques showed that this risk was minor. Second, we used data derived from a RCT instead of a cohort, which potentially limits the generalizability of our results. Third, we evaluated a limited number of predictors in this study and genetic and other biological risk indicators, for example, were not included. This was due to the relatively small population size and our pre-selection criteria for potential predictors: predictors had to be both identified before in multiple studies and easily obtainable in GP practice . Finally, in this study, the use of the PHQ-9 with a cut-off score of 10 or more rendered a higher cumulative incidence of depression than the MINI. This could be explained by the fact that the PHQ-9 was measured more frequently than the MINI. Also, the PHQ-9 was self-reported instead of administered with a diagnostic interview by a trained research assistant. However, it is possible that depression was sometimes over-diagnosed with the PHQ-9 due to potential overlap of (somatic) symptoms of the chronic disease and those of depression[55].

In a previous publication we have hypothesized the causes for the lack of effect of the Step-Dep intervention as compared to care as usual in preventing incident MDD at 12 months of follow-up[20], which we assume also explain the lack of effect at 24 months of follow-up. In summary, a first explanation could be that subthreshold depression was potentially over-diagnosed in our population, whereas stepped-care may be more effective in patients with more severe symptoms[56]. Secondly, fewer patients than expected were treated with the more intensive treatment steps. This was partly caused by the fact that a considerable proportion of patients did not want to start one or more of the treatment steps. This may indicate that our program did not sufficiently match their need for care. Furthermore, this was in part due to the low PHQ-9 scores of 6.7 on average at three months after baseline measurements, which made only a relatively small proportion of the patients eligible for more intensive treatment steps. The drop in PHQ-9 scores between baseline and three months of follow-up in both groups exceeded the expectations of spontaneous recovery alone[57]. It is unlikely that either of the groups received any specific treatment during this period. The Step-Dep program entailed an initial period of watchful waiting and Dutch primary care clinical guidelines recommend a similar waiting period before starting treatment for subthreshold depression[58]. Additionally, screening for depression alone does not change the management of depression in primary care[59]. We argue that the decrease in depressive symptoms may partly be caused by attention, regression to the mean, or patients' self-insight into their mental symptoms and problems. Finally, depressive and anxiety symptoms slightly improved over time in both groups, possibly indicating that usual care is already of reasonable quality and, therefore, the room for improvement for new interventions over usual care may be limited.

Our multivariable model consisted of four predictors of MDD incidence. Firstly, baseline depression severity level is the most frequently found and often strongest predictor of incident depression in other studies in patients with DM2[21,22,25] or CHD[28,31,32]. In line with these findings, in our model a clinically relevant baseline difference in depressive symptoms of five points on the PHQ-9, translated to an almost five times increased risk of developing a MDD during two years. This factor was used as a continuous variable in which the severity level predicts the occurrence of a depressive episode, which supports the concept of a gradual risk of depression. Secondly, the anxiety level at baseline was an important predictor of MDD. Anxiety has been frequently appointed as an important risk factor for depression in DM2[21] and CHD populations[30,31]. Predictors are not necessarily

etiological factors[60]. Nonetheless, as anxiety is also known for its high comorbidity with depression, the assumption that reducing anxiety will have a positive effect on depressive symptoms and MDD incidence seems defendable. Thirdly, the risk the occurrence of stressful life-events pose, has been demonstrated before in patients with CHD[28]. Although most of our knowledge on the role of stressful life-events as predictors of depression cover a short period of time[61], more recent research has shown their long-term risk[62]. This would imply that healthcare providers should not only be temporarily alert on the negative influence on mental health of stressful life-events, but should also be aware of deferred effects. Fourthly, the presence of more than three chronic diseases was identified as a predictor of MDD in our study, in concordance with results in a DM2 population of Fisher et al.[24] Interestingly, the presence of either DM2, CHD or both was not a predictor in our study, which suggests that these patients are at the same risk of incident depression. As all included patients in Fisher's and our study had at least one chronic disease, a discrimination between the predictive values of no chronic disease versus only one versus multiple chronic diseases could not be made. The specific importance of an increased number of diseases as opposed to the risk of a chronic disease has also been demonstrated previously in a primary care population with subthreshold depression[63] and several elderly populations[64]. Why the number of diseases would matter in itself, can perhaps be understood from findings from qualitative interviews. Step-Dep patients explained that chronic diseases indirectly lead to depression, as they diminish future perspectives and cause disability[40], which might be subjective to a certain "threshold" burden of disease. Finally, in contrast to findings in multiple other studies, female sex[24,27,29,31]and a history of depression[25,27,29] did not predict incident MDD in our study. These factors were also not univariately associated with incident depression in our data. A history of depression was self-reported in our study. Perhaps patients over-reported this, as it was not required that they received treatment for this depressive episode, which might explain the lack of an univariate correlation with incident depression.

The model rendered in this study had good discriminative properties with an AUC of 0.80 with the use of only four predictors that are relatively easily obtained by the GP. This makes this prediction model practically viable. It could assist as a tool to both improve the (early) recognition of depression in primary care patients with DM2 and/or CHD and indicate which patients need further care. As chronic care in the Netherlands is being delegated more and more to primary care practice nurses, such a tool might prove useful in their and the GPs' regular check-ups. In

practice, this would not only entail that in patients with DM2 and/or CHD, GPs and practice nurses standardly inquire about symptoms of depression and anxiety during regular checkups, but also that in those with multiple chronic diseases next to their DM2 or CHD, who suffered a recent stressful life-event, the presence and course of depressive and anxiety symptoms are assessed and monitored over time with, for example, the PHQ-9 and HADS. Whenever depressive or anxiety symptoms are clinically severely elevated or significantly deteriorate over time, treatment should be offered according to the patients' need for care. By reducing both depressive and anxiety symptoms, perhaps MDD and its negative consequences can be averted.

Future research should focus on the external validation to test the generalizability of our results, for example on DM2 and/or CHD patients without subthreshold depression, or outside the Dutch setting. Subsequently, studies are required to investigate the influence of the prediction model on decision making and patient outcomes. Consecutively, future research should evaluate whether the suggested enhanced vigilance strategies in combination with depression prevention programs that only target those with all four indicated predictors present and aim to reduce both anxiety and depressive symptoms, are cost-effective[65].

REFERENCES

1. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 2006;3:2011–30.

2. WHO. World Health Statistics 2017: Monitoring health for the SDGs. 2017.

3. Roy T, Lloyd CE. Epidemiology of depression and diabetes: A systematic review. J. Affect. Disord.; 2012;142:S8–21.

4. Rudisch B, Nemeroff CB. Epidemiology of comorbid coronary artery disease and depression. Biol. Psychiatry. 2003;54:227–40.

5. Lin EHB, Katon W, Von Korff M, Rutter C, Simon GE, Oliver M, et al. Relationship of depression and diabetes self-care, medication adherence, and preventive care. Diabetes Care. 2004;27:2154–60.

6. Gehi A, Haas D, Pipkin S. Depression and medication adherence in outpatients with coronary heart disease. Arch Intern Med. 2005;165:2508–13.

7. Ali S. The association between depression and health-related quality of life in people with type 2 diabetes: a systematic literature review. Diabetes. Metab. Res. Rev. 2010;26:75–89.

8. Rutledge T, Bittner V, Olson MB, Linke SE, Cornell CE, Eteiba W, et al. Depression and Cardiovascular Health Care Costs Among Women With Suspected Myocardial Ischemia (Women's Ischemia Syndrome Evaluation) Study. Jac. American College of Cardiology Foundation; 2009;53:176–83.

9. Bosmans JE, Adriaanse MC. Outpatient costs in pharmaceutically treated diabetes patients with and without a diagnosis of depression in a Dutch primary care setting. BMC Health Serv. Res. 2012;12:46.

10. Katon W, Lin EHB, Von Korff M, Ciechanowski P, Ludman E, Young B, et al. Integrating depression and chronic disease care among patients with diabetes and/or coronary heart disease: The design of the TEAMcare study. Contemp. Clin. Trials.; 2010;31:312–22.

11. Sullivan M, O'Connor P, Feeney P, Hire D, Simmons DL, Raisch D, et al. Depression Predicts All-Cause Mortality. Diabetes Care. 2012;35:1708–15.

12. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. Lancet; 2009;374:609–19.

13. National Collaborating Centre for Mental Health. Depression in adults with a chronic physical health problem. The NICE Guideline of Treatment and Management . 2010.

14. Chisholm D, Sanderson K, Ayuso-Mateos JL, Saxena S. Reducing the global burden of depression: Population-level analysis of intervention cost-effectiveness in 14 world regions. Br. J. Psychiatry. 2004;184:393–403.

15. Cuijpers P, van Straten A, Smit F, Mihalopoulos C, Beekman A. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. Am. J. Psychiatry. 2008;165:1272–80.

16. van Zoonen K, Buntrock C, Ebert DD, Smit F, Reynolds CF, Beekman ATF, et al. Preventing the onset of major depressive disorder: A meta-analytic review of psychological interventions. Int. J. Epidemiol. 2014;43:318–29.

17. Bower P, Gilbody S. Stepped care in psychological therapies: Access, effectiveness and efficiency. Narrative literature review. Br. J. Psychiatry. 2005;186:11–7.

18. van Dijk SEM, Pols AD, Adriaanse MC, Bosmans JE, Elders PJM, van Marwijk HWJ, et al. Cost-effectiveness of a stepped-care intervention to prevent major depression in patients with type 2 diabetes mellitus and/or coronary heart disease and subthreshold depression: design of a cluster-randomized controlled trial. BMC Psychiatry. 2013;13:128.

19. Davidson SK, Harris MG, Dowrick CF, Wachtler CA, Pirkis J, Gunn JM. Mental health interventions and future major depression among primary care patients with subthreshold depression. J. Affect. Disord.; 2015;177:65–73.

20. Pols AD, Van Dijk SE, Bosmans JE, Hoekstra T, Van Marwijk HWJ, Van Tulder MW, et al. Effectiveness of a stepped-care intervention to prevent major depression in patients with type 2 diabetes mellitus and/or coronary heart disease and subthreshold depression: A pragmatic cluster randomized controlled trial. PLoS One. 2017;12.

21. Bot M, Pouwer F, Ormel J, Slaets JPJ, de Jonge P. Predictors of incident major depression in diabetic outpatients with subthreshold depression. Diabet. Med. 2010;27:1295–301.

22. Pibernik-Okanovic M, Begic D, Peros K, Szabo S, Metelko Z. Psychosocial factors contributing to persistent depressive symptoms in type 2 diabetic patients: a Croatian survey from the European Depression in Diabetes Research Consortium. J. Diabetes Complications. 2008;22:246–53.

23. Badawi G, Pagé V, Smith KJ, Gariépy G, Malla A, Wang J, et al. Self-rated health: A predictor for the three year incidence of major depression in individuals with Type II diabetes. J. Affect. Disord. 2013;145:100–5.

24. Fisher L, Skaff MM, Mullan JT, Arean P, Glasgow R, Masharani U. A longitudinal study of affective and anxiety disorders, depressive affect and diabetes distress in adults with type 2 diabetes. Diabet. Med. 2008;25:1096–101.

25. Katon W, Ph JR, M EHBLMD, D SRHM, M PCMD, Evette J Ludman Ph, et al. Depression and Diabetes: Factors Associated With Major Depression at Five-Year Follow-Up. Psychosomatics. 2011;50:570–9.

26. Bot M, Pouwer F, Ormel J, Slaets JPJ, de Jonge P. Predictors of incident major depression in diabetic outpatients with subthreshold depression. Diabet. Med. England; 2010;27:1295–301.

27. Nefs G, Pouwer F, Denollet J, Pop V. The course of depressive symptoms in primary care patients with type 2 diabetes: Results from the Diabetes, Depression, Type D Personality Zuidoost-Brabant (DiaDDZoB) Study. Diabetologia. 2012;55:608–16.

28. Doyle F, McGee H, Delaney M, Motterlini N, Conroy R. Depressive vulnerabilities predict depression status and trajectories of depression over 1 year in persons with acute coronary syndrome. Gen. Hosp. Psychiatry.; 2011;33:224–31.

29. Spijkerman TA, Van Den Brink RHS, Jansen JHC, Crijns HJGM, Ormel J. Who is at risk of post-MI depressive symptoms? J. Psychosom. Res. 2005;58:425–32.

30. Pedersen SS, Denollet J, van Gestel YRBM, Serruys PW, van Domburg RT. Clustering of psychosocial risk factors enhances the risk of depressive symptoms 12-months post percutaneous coronary intervention. Eur. J. Cardiovasc.

Prev. Rehabil. 2008;15:203–9.

31. Ossola P, Paglia F, Pelosi A, De Panfilis C, Conte G, Tonna M, et al. Risk factors for incident depression in patients at first acute coronary syndrome. Psychiatry Res. 2015;228:448–53.

32. Kang H, Stewart R, Bae K, Kim S, Shin I, Hong Y, et al. Predictors of depressive disorder following acute coronary syndrome: Results from K-DEPACS and EsDEPACS. J. Affect. Disord.; 2015;181:1–8.

33. Kroenke K SR. The PHQ-9: a new depression diagnostic and severity measure. Psychiatr Ann. 2002;32:509–515.

34. Lamers F, Jonkers CCM, Bosma H, Penninx BWJH, Knottnerus JA, van Eijk JTM. Summed score of the Patient Health Questionnaire-9 was a reliable and valid method for depression screening in chronically ill elderly patients. J. Clin. Epidemiol. 2008;61:679–87.

35. Sheehan DV, Lecrubier Y SK. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatr. 1998;59:22–33.

36. Van Vliet IM, De Beurs E. Het Mini Internationaal Neuropsychiatrisch Interview (MINI): Een kort gestructureerd diagnostisch psychiatrisch interview voor DSM-IV-en ICD-10-stoornissen. Tijdschr. Psychiatr. 2007;49:393–7.

37. Meader N, Mitchell AJ, Chew-graham C, Goldberg D, Rizzo M, Bird V, et al. Case identification of depression in patients with chronic physical health problems: a diagnostic accuracy meta-analysis of 113 studies. 2011;808–20.

38. Moriarty AS, Gilbody S, McMillan D, Manea L. Screening and case finding for major depressive disorder using the Patient Health Questionnaire (PHQ-9): A meta-analysis. Gen. Hosp. Psychiatry; 2015;37:567–76.

39. van der Zwaan GL, van Dijk SEM, Adriaanse MC, van Marwijk HWJ, van Tulder MW, Pols AD, et al. Diagnostic accuracy of the Patient Health Questionnaire-9 for assessment of depression in type II diabetes mellitus and/or coronary heart disease in primary care. J. Affect. Disord. 2015;190:68–74.

40. Pols A, Schipper K, Overkamp D, van Marwijk H, van Tulder M, Adriaanse M. Patients' and practice nurses' perceptions of depression in patients with diabetes type 2 and/or coronary heart disease screened for

subthreshold depression. submitted.

41. Spinhoven P, Ormel J, Sloekers PP, Kempen GI, Speckens AE, Van Hemert AM. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. Psychol Med. 1997;27:363–70.

42. Robins LN, Helzer JE, Croughan J, Ratcliff KS. National Institute of Mental Health Diagnostic Interview Schedule. Arch Gen Psychiatry. 1981;38:381–9.

43. Kriegsman DMW, Penninx BWJH, Van Eijk JTM, Boeke a. JP, Deeg DJH. Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly. A study on the accuracy of patients' self-reports and on determinants of inaccuracy. J. Clin. Epidemiol. 1996;49:1407–17.

44. Twisk J. Different Methods to Analyse the Results of a Randomized Controlled Trial with More Than One Follow-Up Measurement. In: K. van Montfoort, J. Oud & WG, editor. Dev. Stat. Eval. Clin. Trials.; 2014. p. 177–93.

45. McCulloch CE NJ. Generalized Linear Mixed Models. Encyclopedia of Biostatistics. John Wiley & Sons; 2005.

46. Twisk JW. Applied longitudinal data analysis for epidemiology: a practical guide. Cambridge University Press; 2013.

47. Ali S, Stone MA, Peters JL, Davies MJ, Khunti K. The prevalence of co-morbid depression in adults with Type 2 diabetes: a systematic review and meta-analysis. Diabet. Med. 2006;23:1165–73.

48. Comijs HC, Nieuwesteeg J, Kok R, van Marwijk HW, van der Mast RC, Naarding P, et al. The two-year course of late-life depression; results from the Netherlands study of depression in older persons. BMC Psychiatry.; 2015;15:1.

49. de Boer MR, Waterlander WE, Kuijper L, Steenhuis I, Twisk J. Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate. Int. J. Behav. Nutr. Phys. Act. 2015;12:4.

50. Van Buuren, S; Groothuis-oudshoorn K. mice : Multivariate Imputation by Chained. J. Stat. Softw. 2011;45.

51. Rubin DB. Multiple imputation for nonresponse in surveys. John Wiley & Sons; 2004.

52. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. Eur. Urol. 2015;67:1142–51.

53. Heymans MW, Buuren S Van, Knol DL, Van W, Vet HCW De. Variable selection under multiple imputation using the bootstrap in a prognostic study. 2007;10:1–10.

54. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. BMC Med. Res. Methodol. 2009;9:57.

55. Reddy P, Philpot B, Ford D, Dunbar JA. Identification of depression in diabetes: the efficacy of PHQ-9 and HADS-D. Br. J. Gen. Pract. England; 2010;60:e239-45.

56. Cuijpers P, Koole SL, van Dijke A, Roca M, Li J, Reynolds CF. Psychotherapy for subclinical depression: meta-analysis. Br. J. psychiatry. 2014;205:268–74.

57. van't Veer-Tazelaar PJ, van Marwijk HWJ, van Oppen P, van Hout HPJ, van der Horst HE, Cuijpers P, et al. Stepped-care prevention of anxiety and depression in late life: a randomized controlled trial. Arch. Gen. Psychiatry. 2009;66:297–304.

58. Depressie N. M44 NHG-Standaard Depressie. Huisarts&Wetenschap. 2012;55:252–9.

59. Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: A meta-analysis. Cmaj. 2008;178:997–1003.

60. Moons K, Royston P, Vergouwe Y, Grobbee D, Altman D. Prognosis and prognostic research: what, why, and how? BMJ. 2009;338.

61. Kendler KS, Karkowski LM, Prescott CA. Stressful Life Events and Major Depression: Risk Period, Long-Term Contextual Threat, and Diagnostic Specificity. J. Nerv. Ment. Dis. 1998;186:661–9.

62. Assari S, Lankarani MM. Stressful life events and risk of depression 25 years later: Race and gender differences. Front. Public Heal. 2016;4:49.

63. Cuijpers P, Smit F, Willemse G. Predicting the onset of major depression in subjects with subthreshold depression in primary care: A prospective study. Acta Psychiatr. Scand. 2005;111:133–8.

64. Vink D, Aartsen MJ, Schoevers RA. Risk factors for anxiety and depression in the elderly: A review. J. Affect. Disord. 2008;106:29–44.

65. Steyerberg E, Moons KGM, van der Windt D, Hayden J, Perel P, Schroter S, et al. Prognosis research strategy PROGRESS series 3: prognostic model research. PLoS Med. 2013;10:e1001381.

Table 1 Patients' baseline characteristics at baseline in intervention group, care as usual group and total sample

Characteristics	Total sample (N=236)	Intervention (N=96)	Care as usual (N=140)
Female	107/236 (45.3)	42/96 (43.8)	65/140 (46.4)
Age, mean (SD)	67.5 (10.0)	67.8 (9.2)	67.3 (10.5)
Stressful life-event	112/210 (53.3)	48/89 (53.9)	64/121 (52.9)
Positive history of depression	113/210 (53.8)	54/89 (60.7)	59/121 (48.8)
ICPC diagnosis DM2 and/or CHD			
Diabetes Mellitus type 2 (DM2)	88/236 (37.3)	38/96 (39.6)	50/140 (35.7)
Coronary Heart Disease (CHD)	86/236 (36.4)	36/96 (37.5)	50/140 (35.7)
DM2 and CHD	62/236 (26.3)	22/96 (22.9)	40/140 (28.6)
More than 3 chronic diseases	98/210 (46.7)	38/89 (42.7)	60/121 (49.6)
PHQ-9 at baseline, mean (SD)	9.4 (3.2)	9.5 (3.1)	9.3 (3.2)
Anxiety HADS, mean (SD)	6.5 (3.8)	6.9 (3.7)	6.3 (3.9)
Depression HADS, mean (SD)	6.5 (3.8)	6.9 (3.9)	6.1 (3.7)
Marital status			
Married/living together	122/220 (55.5)	55 (61.1)	67/130 (51.5)
Single/divorced/widowed	98/220 (44.5)	35 (38.9)	63/130 (48.5)
Both parents born in the Netherlands	186/220 (84.5)	74/90 (82.2)	112/130 (86.2)
Rural residential area	99/236 (41.9)	42 (43.8)	57/140 (40.7)
Unemployed/sick	26/220 (11.8)	12/90 (13.3)	14/130 (10.8)
Level of education			
Low	89/220 (40.5)	33/90 (36.7)	56/130 (43.1)
Average	60/220 (27.3)	22/90 (24.4)	38/130 (29.2)
High	71/220 (32.3)	35/90 (38.9)	36/130 (27.7)
Current smoker	39/219 (17.8)	16/90 (17.8)	23/129 (17.8)
Alcohol use above norm	63/219 (28.8)	29/90 (32.2)	34/129(26.4)
Exercise under norm	141/219 (64.4)	56/90 (62.2)	85/129 (65.9)
BMI, mean (SD)	28.9 (6.1)	29.4 (6.8)	28.5 (5.6)
Locus of Control, mean (SD)	7.9 (4.2)	8.3 (4.2)	7.6 (4.1)
Social support, mean (SD)	36.3 (9.2)	35.8 (9.0)	36.7 (9.5)
Dysthymia	13/236 (5.5)	6/96 (6.3)	7/140 (5.0)
Onset of depression after age of 55	101/210 (48.1)	38/89 (42.7)	63/121 (52.1)

Numbers are percentages unless stated otherwise; Abbreviations: BMI = Body Mass Index; PHQ-9, Patient Health Questionnaire-9; HADS, Hospital Anxiety and Depression Scale; SD, Standard Deviation.

Table 2 Results of the mixed model and GEE long-term effectiveness analyses

Cumulative incidence of depression (n/N) %	Intervention	Care as usual	Corrected analyses*		Crude analyses	
Baseline	0	0	**OR (95%CI)**	**P-value**	**OR (95%CI)**	**P-value**
T6	(5/84) 6.0	(10/125) 8.0	0.82 (0.19; 3.51)	0.79	0.90 (0.32; 2.50)	0.84
T12	(9/82) 11.0	(12/118) 10.2	1.44 (0.46; 4.47)	0.53	1.20 (0.49; 2.92)	0.70
T24	(13/77) 16.9	(17/105) 16.2	1.23 (0.50; 3.02)	0.66	1.11 (0.51; 2.44)	0.79
Overall effect	n.a	n.a	1.37 (0.52;3.55)	0.52	1.11 (0.49;2.49)	0.80
PHQ mean (SD)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	9.53 (3.14)	9.28 (3.23)	**B (95%CI)**	**P-value**	**B (95%CI)**	**P-value**
T3	6.68 (4.55)	6.58 (4.21)	-0.39 (-1.52; 0.74)	0.50	-0.03 (-1.17; 1.11)	0.96
T6	6.10 (4.43)	6.12 (4.41)	-0.37 (-1.50; 0.76)	0.52	-0.17 (-1.30; 0.95)	0.76
T9	6.28 (4.31)	6.46 (4.51)	-0.48 (-1.62; 0.65)	0.40	-0.40 (-1.53; 0.73)	0.49
T12	6.60 (5.23)	6.29 (4.46)	-0.09 (-1.20; 1.02)	0.88	-0.03 (-1.13; 1.07)	0.96
T24	5.81 (4.76)	5.15 (4.33)	0.00 (-1.18; 1.19)	0.88	0.02 (-1.15; 1.19)	0.97
Overall effect	n.a	n.a	0.29 (-1.15; 0.58)	0.52	-0.13 (-0.99; 0.73)	0.77
Perceived recovery (%)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	n.a	n.a	**OR (95%CI)**	**P-value**	**OR (95%CI)**	**P-value**
T3	40.3%	49.5%	0.78 (0.42; 1.45)	0.44	0.64 (0.36; 1.15)	0.14
T6	48.8%	45.5%	1.46 (0.79; 2.69)	0.23	1.15 (0.65; 2.02)	0.64
T9	55.0%	48.7%	1.47 (0.79; 2.75)	0.22	1.30 (0.74; 2.30)	0.91
T12	55.6%	58.1%	1.04 (0.56; 1.92)	0.91	0.91 (0.51; 1.61)	0.74
T24	68.0%	57.1%	2.38 (1.21; 4.67)	0.01	2.04 (1.08; 3.87)	0.03
Overall effect	n.a	n.a	1.32 (0.87; 2.00)	0.19	1.10 (0.75; 1.62)	0.61
HADS-A mean (SD)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	6.91 (3.74)	6.25 (3.90)	**B (95%CI)**	**P-value**	**B (95%CI)**	**P-value**
T3	6.35 (4.04)	6.29 (3.97)	-0.27 (-1.13; 0.60)	0.54	-0.13 (-1.00; 0.74)	0.76
T6	5.70 (4.10)	6.63 (4.00)	-1.04 (-1.91; -0.18)	0.02	-1.04 (-1.91; -0.18)	0.02
T9	6.16 (4.24)	6.03 (4.04)	-0.49 (-1.35; 0.38)	0.27	-0.45 (-1.31; 0.42)	0.31
T12	5.77 (4.69)	5.83 (3.99)	-0.50 (-1.37; 0.38)	0.27	-0.43 (-1.31; 0.44)	0.33
T24	5.45 (4.46)	5.06 (3.90)	-0.59 (-1.50; 0.31)	0.20	-0.48 (-1.38; 0.43)	0.30
Overall effect	n.a	n.a	-0.59 (-1.23; 0.06)	0.08	-0.52 (-1.17; 0.13)	0.12
HADS-D mean (SD)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	6.93 (3.87)	6.11 (3.73)	**B (95%CI)**	**P-value**	**B (95%CI)**	**P-value**
T3	6.14 (4.16)	6.21 (3.87)	-0.26 (-1.12; 0.60)	0.55	-0.29 (-1.15; 0.56)	0.51
T6	5.82 (3.79)	5.75 (4.03)	-0.22 (-1.07; 0.64)	0.62	-0.32 (-1.18; 0.53)	0.46
T9	6.36 (4.04)	6.07 (4.08)	-0.21 (-1.06; 0.65)	0.63	-0.24 (-1.09; 0.61)	0.58
T12	6.09 (4.20)	6.11 (4.22)	-0.41 (-1.27; 0.46)	0.36	-0.50 (-1.36; 0.36)	0.26
T24	5.59 (4.66)	4.92 (3.90)	-0.41 (-1.30; 0.48)	0.37	-0.48 (-1.37; 0.41)	0.29
Overall effect	n.a	n.a	-0.30 (-0.94; 0.33)	0.35	-0.37 (-1.00; 0.26)	0.25

Abbreviations: 95%CI, 95% Confidence Interval; HADS-A, Hospital Anxiety and Depression Scale-Anxiety; HADS-D, Hospital Anxiety and Depression Scale-Depression; n.a, not applicable; PHQ-9, Patient Health Questionnaire-9;

*Corrected for: baseline values of the outcome, age, gender, marital status, employment status, level of education, co-existence of DM2 and CHD, alcohol use, number of depressive episodes in history and age of onset of depression. The baseline value of the outcome is not added as an extra variable in the corrected analyses of the overall effects since it is already incorporated in the crude overall analyses.

Table 3 Multivariable prediction model of incident depression during two-year follow-up

Predictor	RC	OR	95% CI	P-value
Female sex		-	-	-
Age		-	-	-
Somatic disorder DM2 CHD DM2and CHD		-	-	-
History of depression		-	-	-
Baseline depression scores	0.32 p.p.i.	1.37	1.20; 1.55	0.00
Baseline anxiety scores	0.12 p.p.i.	1.13	1.02; 1.25	0.01
Stressful life-event in past year	0.74	2.10	1.02; 4.32	0.04
>3 chronic illnesses	0.78	2.19	1.12; 4.25	0.02
Randomization status I vs C	0.14	1.15	0.58; 2.29	0.68

RC regression coefficient; p.p.i. per point increase; 95% CI 95% confidence interval; OR odds ratio, an OR > 1 reflects a higher probability the outcome an incident depression and an OR < 1 reflects a lower probability compared with the reference category. OR estimated after multiple imputation (n = 25 datasets) with p-value of 0.157. Linear predictor corrected after bootstrapping = -4.1147 + 0.131* Randomization status + 0.7167* >3 chronic illnesses + 0.680* stressful life-event in past year + 0.1118* baseline anxiety scores + 0.2868* baseline depression scores

SUPPLEMENTARY INFORMATION

Contributors

AP constructed the design of this study, performed all statistical analyses and drafted the manuscript. MCA, MvT, HvM constructed the design of the study and revised the manuscript. JB and SvD constructed the design of the Step-Dep study and revised the manuscript. MH collaborated on the statistical analyses and revised the manuscript. The final manuscript was read and approved by all authors.

Funding

This study is funded by ZonMw, the Netherlands Organisation for Health Research and Development (project number 80-82310-97-12110). The sponsor had no role in the design and conduct of the present study or in the writing of the manuscript.

Competing Interests

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Trial registration and ethical approval

The study was performed in accordance with the declaration of Helsinki (2008) and the Dutch Medical Research involving Human Subjects Act (WMO). The protocol was approved by the medical ethics committee of the VU University Medical Centre (NL39261.029.12, registration number 2012/223), and registered in the Dutch Trial Register (registration number 3715).

Data sharing

Full dataset and statistical code is available from the corresponding author. Consent was not obtained but the presented data are anonymised and risk of identification is low.

Transparency

The corresponding author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; no important aspects of this study have been omitted; and any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Acknowledgements

The authors would like to thank Marcella van der Linden, Lucca Vledder and Mieke Schlattmann for their contribution in the data collection for this study and Jos Twisk for his help in the long-term effectiveness analyses. We also would like to thank all the participating general practices and the research networks of general practitioners (ANH, THOON and LEON) for their participation and collaboration in the implementation and execution of the study. Furthermore, this study has been possible thanks to all Step-Dep participants.

SUPPLEMENTARY FILES

S1 Original protocol

S2 TRIPOD statement checklist

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	4-5
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	5
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	6
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	6
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	6
	5b	D;V	Describe eligibility criteria for participants.	6
	5c	D;V	Give details of treatments received, if relevant.	7
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	6
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	6
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	7
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	6
Sample size	8	D;V	Explain how the study size was arrived at.	Ref protocol
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	8
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	7-8
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	8-9
	10c	V	For validation, describe how the predictions were calculated.	n.a.
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	9
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	n.a.
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	n.a.
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	n.a.
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	10
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	10, table 1
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	n.a.
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	n.a.
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	n.a.
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	Table 3
	15b	D	Explain how to the use the prediction model.	10-11
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	10-11
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	n.a.
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	12
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	n.a.
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	12-13
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	13-14
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	24-26
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	24

TRIPOD Checklist: Prediction Model Development and Validation

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

BMJ Open

Two-year effectiveness of a stepped-care depression prevention intervention and predictors of incident depression in primary care patients with diabetes type 2 and/or coronary heart disease and subthreshold depression; data from the Step-Dep cluster randomized controlled trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020412.R2
Article Type:	Research
Date Submitted by the Author:	01-Jun-2018
Complete List of Authors:	Pols, Alide; VU University, Department of Health Sciences Adriaanse, Marcel; Institute of Health Sciences, Vrije Universiteit Amsterdam Tulder, Maurits; University of Amsterdam, Health Sciences Heymans, Martijn; VU University Medical Center, Department of Epidemiology and Biostatistics; VU University, Department of Health Sciences Section Methodology and Applied Biostatistics Faculty of Earth and Life Sciences Bosmans, J; VU University Amsterdam, Department of Health Sciences van Dijk, Susan; Vrije Universiteit Amsterdam, Department of Health Sciences Van Marwijk, H; University of Brighton, Division of Primary Care and Public Health, Brighton and Sussex Medical School, Mayfield House
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Cardiovascular medicine, Diabetes and endocrinology, Mental health, General practice / Family practice
Keywords:	Depression & mood disorders < PSYCHIATRY, DIABETES & ENDOCRINOLOGY, Coronary heart disease < CARDIOLOGY, PREVENTIVE MEDICINE

Two-year effectiveness of a stepped-care depression prevention intervention and predictors of incident depression in primary care patients with diabetes type 2 and/or coronary heart disease and subthreshold depression; data from the Step-Dep cluster randomized controlled trial

Second revised version June 1st 2018

Alide D Pols*[1,2], Marcel C Adriaanse[1], Maurits W van Tulder[1], Martijn W Heymans[3], Judith E Bosmans[1], Susan E van Dijk[1], Harm W J van Marwijk[2,4]

[1] Department of Health Sciences, Amsterdam Public Health research institute, Faculty of Science, Vrije Universiteit, Amsterdam, The Netherlands

[2] Department of General Practice & Elderly Care Medicine, Amsterdam Public Health research institute, VU University Medical Centre, Amsterdam, The Netherlands

[3] Department of Clinical Epidemiology and Biostatistics VU Medical Centre Amsterdam, The Netherlands

[4] Division of Primary Care and Public Health, Brighton and Sussex Medical School, Mayfield House, University of Brighton, Brighton, United Kingdom

* Corresponding author: a.d.pols@vu.nl

Keywords: major depressive disorder, subthreshold depression, diabetes mellitus type 2, coronary heart disease, effectiveness, stepped care, prediction model

ABSTRACT

Introduction Major depressive disorders (MDD), diabetes mellitus type 2 (DM2) and coronary heart disease (CHD) are leading contributors to the global burden of disease and often co-occur.

Objectives To evaluate the two-year effectiveness of a stepped-care intervention to prevent MDD compared to usual care and to develop a prediction model for incident depression in DM2 and/or CHD patients with subthreshold depression.

Methods Data of 236 Dutch primary care DM2/CHD patients with subthreshold depression (Patient Health Questionnaire 9 (PHQ-9) score ≥6, no current MDD according to the Mini International Neuropsychiatric Interview (DSM-IV criteria)), who participated in the Step-Dep trial were used. A PHQ-9 score of ≥10 at minimally one measurement during follow-up (at 3, 6, 9, 12 and 24 months) was used to determine the cumulative incidence of MDD. Potential demographic and psychological predictors were measured at baseline via web-based self-reported questionnaires and evaluated using a multivariable logistic regression model. Model performance was assessed with the Hosmer–Lemeshow test, Nagelkerke's R^2 explained variance and Area Under the Receiver Operating Characteristic curve (AUC). Bootstrapping techniques were used to internally validate our model.

Results 192 patients (81%) were available at two-year follow-up. The cumulative incidence of MDD was 97/192 (51%). There was no statistically significant overall treatment effect over 24 months of the intervention (OR 1.37; 95% CI 0.52; 3.55). Baseline levels of anxiety, depression, the presence of >3 chronic diseases and stressful life-events predicted the incidence of MDD (AUC 0.80 interquartile range (IQR) 0.79-0.80; Nagelkerke's R^2 0.34 IQR 0.33-0.36).

Conclusion A model with four factors predicted depression incidence during two-year follow-up in patients with DM2/CHD accurately, based on the AUC. The Step-Dep intervention did not influence the incidence of MDD. Future depression prevention programs should target patients with these four predictors present, and aim to reduce both anxiety and depressive symptoms.

TRIAL REGISTRATION NUMBER

Dutch Trial Register NTR3715 http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=3715

STRENGTHS AND LIMITATIONS OF THIS STUDY

- This study provides a prediction model of incident MDD in DM2 and/or CHD patients with subthreshold depression, which could assist healthcare providers in its detection and facilitate targeting indicated prevention to highest risk patients
- Only predictors that are readily available or easily obtained in practice were used in the multivariable model, which enhances the practical use of the model
- This study had a relatively long follow-up and outcomes were frequently measured, whereas drop-out rates were relatively low and missing values imputed
- The relatively small study population might have caused over-optimism of the prediction model, but an internal validation procedure with bootstrapping techniques showed that this risk was minor
- Data were derived from a RCT, but statistically non-significant intervention effects for incident MDD at both 12- and 24-months follow-up justify using the Step-Dep population as a cohort

INTRODUCTION

Depression is a major and increasing contributor to the global burden of disease[1], whereas coronary heart disease (CHD) and diabetes mellitus type 2 (DM2) rank among the leading causes of morbidity and mortality worldwide[2]. Comorbid depression in patients with DM2 and/or CHD is common[3,4] and has detrimental effects on self-care and medication adherence[5,6], quality of life[7], health status and increases healthcare costs[8,9] and mortality[10,11]. Despite its negative impact, many cases of depression go unrecognized in primary care[12], especially in patients with chronic diseases like DM2 and/or CHD[13]. Additionally, about one-third of those recognized and treated does not respond to current approaches, and over half of those who experience a first onset of a major depressive episode will experience one or more recurrences[14].

Given the significant burden of disease of depression, its poor recognition and the limited effect of current treatment options for it, it would be of great value if incident cases could be averted by early detection and preventive treatment of patients at risk ('indicated prevention'). Meta-analyses have shown that preventive psychological interventions can overall reduce the incidence of MDD in comparison to control groups[15,16]. Offering preventive psychological interventions in a stepped-care format could be an efficient approach, as patients start with minimally intensive evidence-based treatments and only those who do not improve adequately, step up to a treatment of higher intensity[17]. Recently, we conducted a randomized controlled trial in which we evaluated whether a pragmatic nurse-led stepped-care program was effective in reducing the incidence of MDD at 12-months of follow-up in comparison with usual care among patients with DM2 and/or CHD and subthreshold depression (Step-Dep)[18]. Subthreshold depression entails clinically relevant depressive symptoms without fulfilling the criteria for MDD and is a known important risk factor for depression[15,19]. We demonstrated that the Step-Dep intervention was not superior to usual care and the overall cumulative incidence of MDD was lower than expected after one year [20]. However, it may be possible that the follow-up period was too restricted to demonstrate the potential health benefits of the stepped-care program over usual care, or the presence of subthreshold depression alone posed a lower than expected prior risk of MDD in our DM2 and/or CHD population.

Identifying additional major risk factors of incident depression in patients with DM2 and/or CHD might facilitate targeting indicated prevention to patients with highest risk, but also potentially aid in its detection. In patients with DM2, several longitudinal studies have been conducted to determine risk factors for comorbid incident depression. However, these studies have rendered heterogeneous results, due to small patient samples (<80 at follow-up)[21,22], analyses of single factors only[23,24], the use of mixed samples of type 1 diabetes and DM2[25], patients with either no MDD at baseline[23,26] or both with and without depression at baseline[22,24,25,27], and differences across community[23,24], primary care[25,27] and secondary care settings[22,26]. In patients with CHD, the only available longitudinal data are derived from studies in patients with acute coronary syndrome followed-up after hospital discharge[28–32]. Predictors that were repeatedly identified in DM2 or CHD studies were: depression severity at baseline[21,22,25,28,31,32], history of depression[25,27,29], female sex[24,27,29,31] and baseline anxiety levels[21,30,31]. However, data of patients with both DM2 and CHD, non-acute CHD or within primary care settings are scarce. The goal of the present study was twofold: (1) to evaluate the two-year effectiveness of a nurse-led stepped-care intervention to prevent MDD as compared to usual care (Step-Dep); and to (2) develop a prediction model for incident depression during two-year follow-up in primary care patients with DM2 and/or CHD and subthreshold depression.

METHODS

Design

Data of the Step-Dep cluster randomized controlled trial were used. Step-Dep was conducted in 27 general practitioner (GP) practices in three regions in the Netherlands (Amsterdam, Leiden, Twente), between January 2013 and November 2016, including recruitment and two years of follow-up. A statistician blinded to the characteristics of the GP practices performed the randomization of GP practices using a computer generated list of random numbers. Randomization was done at the level of the GP practice, which corresponds to the participating practice nurse, to avoid contamination between the treatment groups, and was stratified for size (less or more than 5000 patients). The study was performed in accordance with the declaration of Helsinki (2008) and the Dutch Medical Research involving Human Subjects Act (WMO). The protocol was approved by the medical ethics committee of the VU University Medical Centre (NL39261.029.12, registration number 2012/223), and registered in the Dutch Trial Register (NTR3715 http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=3715) (S1 Original protocol). Further details on the methods and design of the Step-Dep study have been published elsewhere[18].

Patient and Public Involvement

Patients were not involved in determining the design, the recruitment to or conduct of the study. The medical ethics committee of the VU University Medical Centre assessed the burden of the intervention and participation in the study in general as acceptable for patients. The burden of and satisfaction with the intervention were assessed in a process evaluation with 15 patients. All patients are thanked in the acknowledgements section. Results of the study will be disseminated by letter to all participants.

Patients

Included patients were aged 18 years or more who had an International Classification of Primary Care (ICPC) diagnosis of DM2 and/or CHD and had subthreshold depression identified by screening. Patients with a Patient Health Questionnaire 9 (PHQ-9; range 0-27 with higher scores indicating more severe depressive symptoms) score of six or higher[33,34], and no major depressive disorder according to the Mini International Neuropsychiatric

Interview (MINI)[35,36], were considered to have subthreshold depression. Exclusion criteria were cognitive impairment, psychotic illnesses, a terminal illness, the use of anti-depressant medication, a history of suicide attempt(s), loss of significant other in the past six months, visual impairment, current pregnancy, bipolar disorder, borderline personality disorder or any difficulties completing written questionnaires or visiting the primary care center. A total of 236 patients gave informed consent to participate.

Outcome measure

The outcome measure used was an incident depression (yes/no) defined as a PHQ-9 score of ≥10 at minimally one moment during follow-up (measured at 3, 6, 9, 12 and 24 months after baseline). The PHQ-9 is a widely used and validated instrument that performs well in patients with chronic medical illnesses both as dichotomous diagnosis of major and minor depression and a continuous severity score[34,37]. A cut-off of ≥10 has been shown to be the optimum cutoff for major depression[38], also in this patient group [39]. PHQ-9 was self-reported with web-based or written questionnaires. When these web-based or written questionnaires were not completed, the PHQ-9 was administered by telephone by trained research-assistants, blinded to randomization status.

Potential predictors

The selection of the potential predictors was based on a thorough literature search. Predictors of incident depression that were identified in multiple studies in patients with DM2 or CHD and are routinely available or easily obtained in daily GP practice were used. Additionally, we chose the presence of multiple chronic diseases[24] and stressful life-events[28] although they were identified in single studies only, as these were also indicated as causes of depression by patients and practice nurses in semi-structured interviews as part of the process evaluation of Step-Dep[40], and age[23].

Apart from GP information system derived data on *sex, age* and *ICPC diagnosis of DM2 and/or CHD*, demographics and psychological factors were measured at baseline via web based (or written if preferred) self-reported questionnaires. To take possible effects of the intervention into account, we included randomization status in the selection models as well. Patients in the intervention arm were offered a stepped care prevention program, and

patients in the control arm received care as usual during one year. The stepped care intervention consisted of four sequential but flexible treatment steps, each lasting three months; 1) watchful waiting, 2) guided self-help, 3) problem solving treatment and 4) referral to a general practitioner. After each step, patients with a persisting PHQ-9 score of six or more were offered the next treatment step of the intervention. *Baseline depression levels* were measured with the PHQ-9[33,34]. *Baseline anxiety levels* were measured with the Hospital Anxiety and Depression Scale Anxiety (HADS-A; range 0-21 with higher scores indicating more severe anxiety)[41]. *History of depression* and s*tressful life-events were* self-reported using a subset of the Diagnostic Interview Schedule (DIS)[42]. *Number of co-morbid chronic illnesses* was measured using the self-reported Dutch Questionnaire Chronic Illnesses[43]. This was dichotomized using the median in our sample: three or less versus more than three chronic diseases.

Statistical analyses

The two-year effectiveness of the intervention on the primary and secondary outcomes was analyzed according to the intention to treat principle. Generalised Estimating Equations (GEE) were used for binary outcome variables, and linear mixed models for longitudinal data were used for continuous outcome variables[44]. For each outcome an overall effect over time and separate effects at different time points were estimated by taking time into account as a categorical variable (with five categories: 0-3 months, 3-6 months, 6-9 months, 9-12 months and 12-24 months of follow-up)[45,46]. The main analyses consisted of fully corrected models that were corrected for baseline values of the respective outcome and additionally included the covariates gender[47], age[48], and any other possible confounding variable on which the treatment groups differed at baseline (marital status, employment status, level of education, co-existence of DM2 and CHD, alcohol use, number of depressive episodes in history and age of onset of depression), based on absolute baseline differences judged by the researchers rather than statistical testing[49]. For these analyses, STATA version 14 was used.

Missing data were imputed using multiple imputation according to the Multivariate Imputation by Chained Equations (MICE) algorithm[50] in SPSS version 23. For the imputations, missing at random (MAR) was assumed. Variables that were associated with missing data and variables that were associated with the outcome, were identified and included in the imputation model. Also, all variables in the analysis model (potential predictors and

outcome) were included. The number of imputed datasets was 25 based on the proportion of cases with incomplete measurements; 24%. The subsequent analyses were performed on pooled data according to Rubin's rules[51].

Prediction model

We created a multivariable logistic regression model in SPSS 23 from the baseline variables estimating the probability of having at least one major depression (PHQ ≥10) during the two-year assessment. To calculate the number of potential predictors for developing the prediction model, we used the criterion of 10 events per variable. Continuous variables were checked for linearity with the outcome using spline regression curves and linearity was confirmed. All variables were entered into the logistic model and tested for statistical significance in the presence of the total set of predictors. Individually, the least significant predictor (P-value>0.157, as recommended in the TRIPOD statement (S2 TRIPOD statement checklist), [52], Wald statistic was removed, and the model was refit (backward selection). Randomization status was maintained in the model. This was repeated until we reached a statistical model that only included statistically significant predictors. This was repeated with p-values of 0.05. We also compared the results with complete case analysis (CCA), i.e., all patients with missing data were excluded from the analyses.

We checked the performance of the model with regard to the goodness of fit (Hosmer–Lemeshow test), the explained variation and the discriminative ability of the model. The Nagelkerke's R^2 explained variation is the extent to which the outcome can be predicted by the predictors in the model in current datasets. The discriminative ability is reflected by the area under the receiver operating characteristics curve (AUC). Bootstrapping techniques were used to internally validate our model, i.e., to simulate the performance with respect to the explained variance and the AUC in comparable patient datasets[53]. After that, we calculated the linear predictor of the bootstrapped model with an adjusted intercept and regression coefficients corrected for the shrinkage factor. Performance measures were assessed in each imputed dataset and results were summarized using median values [54]. All analyses were done with SPSS version 23.0 and R software.

RESULTS

Participants

The baseline characteristics of the study population are presented in Table 1. Of the 236 patients included in Step-Dep, 192 patients (81%) completed two years of follow-up. A flowchart of participants through the first 12 months of the Step-Dep study has been published elsewhere[20]. At 24 months of follow up 18 additional patients dropped out (two for unknown motives, seven due to time considerations, four were deceased, three too frail, two unable to contact). We compared the baseline characteristics of patients with missing data to those without. Patients with missing data were more often living alone (61% vs 41%), but no other differences between these groups were found.

There was no statistically significant overall treatment effect over 24 months of the intervention (OR 1.37; 95% CI 0.52 to 3.55), nor at any of the time-points. There were no significant differences in PHQ-9 scores between the study groups at any time-point and the course of PHQ-9 scores over time did not differ significantly between the groups. Results are shown in Table 2. The statistically non-significant intervention effects for incident MDD at both 12-months[20] and 24-months of follow-up justify using the Step-Dep population as a cohort.

Prediction model

The cumulative incidence during two-year follow-up was 97/192 (51%). The multivariable models using p=0.05 and p=0.157[52] were identical. The final model consisted of four predictors: level of anxiety, level of depression, presence of more than three chronic diseases and having suffered a stressful life-event in the past year. This model performed well (Hosmer–Lemeshow test p=0.12 and median of pooled Nagelkerke's R^2 explained variance 0.34 interquartile range (IQR) 0.33-0.36) with good discriminative properties (median of the pooled AUC 0.80 IQR 0.79-0.80). In a CCA with p=0.05, the same predictors remained. In a CCA using p=0.157 [52], the categorical variable DM2/CHD/both also remained.

The risk of an incident MDD during two years of follow-up more than doubled when either more than three chronic diseases were present or a patient had suffered a stressful life-event in the past year. Both higher

depression and anxiety levels at baseline increased the risk of MDD with each incremental point on the PHQ-9 of HADS scales respectively. One point higher on the PHQ-9 at baseline, resulted in a 1.37 higher risk of developing MDD during two years, compared to 1.13 for increasing anxiety levels. With regard to the internal validation of the model, the calibration slope (or shrinkage factor to correct regression coefficients of the original model) was 0.92 IQR 0.91-0.92, the median explained variance was 31% IQR 0.29-0.32 and the AUC 0.78 IQR 0.77-0.78. This means that after corrections for over-optimism, both the performance and discriminative properties of the model remained good. Results are shown in Table 3.

DISCUSSION

This study showed that the Step-Dep intervention was not more effective than usual care in the prevention of MDD at two years of follow-up. The risk of incident MDD during two years of follow-up among patients with DM2 and/or CHD and subthreshold depression, was increased by higher baseline levels of anxiety and depression, the presence of more than three chronic diseases and having suffered a stressful life-event in the past year. This risk was not influenced by a stepped-care intervention aimed at preventing MDD.

Our findings have to be viewed in the context of strengths and limitations of this study. Strengths are its relatively long follow-up with frequent outcome measurements and low drop-out rates. In addition, missing values were imputed using multiple imputation techniques. We only used predictors that are readily available or easily obtained in practice, which enhances the practical use of the model in primary care consultations. Furthermore, testing a multivariable model instead of single factors appointed only the most relevant predictors, which rendered a simple model that is manageable in its use. There were limitations to this study. First, the study population was relatively small, which might have caused over-optimism of the prediction model. This means that it predicts the outcome better in the sample used to develop the model than in new samples, potentially restricting its external validity. However, an internal validation procedure with bootstrapping techniques showed that this risk was minor. Second, we used data derived from a RCT instead of a cohort, which potentially limits the generalizability of our results. Third, we evaluated a limited number of predictors in this study and genetic and other biological risk indicators, for example, were not included. This was due to the relatively small population size and our pre-selection criteria for potential predictors: predictors had to be both identified before in multiple studies and easily obtainable in GP practice . Finally, in this study, the use of the PHQ-9 with a cut-off score of 10 or more rendered a higher cumulative incidence of depression than the MINI. This could be explained by the fact that the PHQ-9 was measured more frequently than the MINI. Also, the PHQ-9 was self-reported instead of administered with a diagnostic interview by a trained research assistant. However, it is possible that depression was sometimes over-diagnosed with the PHQ-9 due to potential overlap of (somatic) symptoms of the chronic disease and those of depression[55].

In a previous publication we have hypothesized the causes for the lack of effect of the Step-Dep intervention as compared to care as usual in preventing incident MDD at 12 months of follow-up[20], which we assume also explain the lack of effect at 24 months of follow-up. In summary, a first explanation could be that subthreshold depression was potentially over-diagnosed in our population, whereas stepped-care may be more effective in patients with more severe symptoms[56]. Secondly, fewer patients than expected were treated with the more intensive treatment steps. This was partly caused by the fact that a considerable proportion of patients did not want to start one or more of the treatment steps. This may indicate that our program did not sufficiently match their need for care. Furthermore, this was in part due to the low PHQ-9 scores of 6.7 on average at three months after baseline measurements, which made only a relatively small proportion of the patients eligible for more intensive treatment steps. The drop in PHQ-9 scores between baseline and three months of follow-up in both groups exceeded the expectations of spontaneous recovery alone[57]. It is unlikely that either of the groups received any specific treatment during this period. The Step-Dep program entailed an initial period of watchful waiting and Dutch primary care clinical guidelines recommend a similar waiting period before starting treatment for subthreshold depression[58]. Additionally, screening for depression alone does not change the management of depression in primary care[59]. We argue that the decrease in depressive symptoms may partly be caused by attention, regression to the mean, or patients' self-insight into their mental symptoms and problems. Finally, depressive and anxiety symptoms slightly improved over time in both groups, possibly indicating that usual care is already of reasonable quality and, therefore, the room for improvement for new interventions over usual care may be limited.

Our multivariable model consisted of four predictors of MDD incidence. Firstly, baseline depression severity level is the most frequently found and often strongest predictor of incident depression in other studies in patients with DM2[21,22,25] or CHD[28,31,32]. In line with these findings, in our model a clinically relevant baseline difference in depressive symptoms of five points on the PHQ-9, translated to an almost five times increased risk of developing a MDD during two years. This factor was used as a continuous variable in which the severity level predicts the occurrence of a depressive episode, which supports the concept of a gradual risk of depression. Secondly, the anxiety level at baseline was an important predictor of MDD. Anxiety has been frequently appointed as an important risk factor for depression in DM2[21] and CHD populations[30,31]. Predictors are not necessarily

etiological factors[60]. Nonetheless, as anxiety is also known for its high comorbidity with depression, the assumption that reducing anxiety will have a positive effect on depressive symptoms and MDD incidence seems defendable. Thirdly, the risk the occurrence of stressful life-events pose, has been demonstrated before in patients with CHD[28]. Although most of our knowledge on the role of stressful life-events as predictors of depression cover a short period of time[61], more recent research has shown their long-term risk[62]. This would imply that healthcare providers should not only be temporarily alert on the negative influence on mental health of stressful life-events, but should also be aware of deferred effects. Fourthly, the presence of more than three chronic diseases was identified as a predictor of MDD in our study, in concordance with results in a DM2 population of Fisher et al.[24] Interestingly, the presence of either DM2, CHD or both was not a predictor in our study, which suggests that these patients are at the same risk of incident depression. As all included patients in Fisher's and our study had at least one chronic disease, a discrimination between the predictive values of no chronic disease versus only one versus multiple chronic diseases could not be made. The specific importance of an increased number of diseases as opposed to the risk of a chronic disease has also been demonstrated previously in a primary care population with subthreshold depression[63] and several elderly populations[64]. Why the number of diseases would matter in itself, can perhaps be understood from findings from qualitative interviews. Step-Dep patients explained that chronic diseases indirectly lead to depression, as they diminish future perspectives and cause disability[40], which might be subjective to a certain "threshold" burden of disease. Finally, in contrast to findings in multiple other studies, female sex[24,27,29,31]and a history of depression[25,27,29] did not predict incident MDD in our study. These factors were also not univariately associated with incident depression in our data. A history of depression was self-reported in our study. Perhaps patients over-reported this, as it was not required that they received treatment for this depressive episode, which might explain the lack of an univariate correlation with incident depression.

The model rendered in this study had good discriminative properties with an AUC of 0.80 with the use of only four predictors that are relatively easily obtained by the GP. This makes this prediction model practically viable. It could assist as a tool to both improve the (early) recognition of depression in primary care patients with DM2 and/or CHD and indicate which patients need further care. As chronic care in the Netherlands is being delegated more and more to primary care practice nurses, such a tool might prove useful in their and the GPs' regular check-ups. In

practice, this would not only entail that in patients with DM2 and/or CHD, GPs and practice nurses standardly inquire about symptoms of depression and anxiety during regular checkups, but also that in those with multiple chronic diseases next to their DM2 or CHD, who suffered a recent stressful life-event, the presence and course of depressive and anxiety symptoms are assessed and monitored over time with, for example, the PHQ-9 and HADS. Whenever depressive or anxiety symptoms are clinically severely elevated or significantly deteriorate over time, treatment should be offered according to the patients' need for care. By reducing both depressive and anxiety symptoms, perhaps MDD and its negative consequences can be averted.

Future research should focus on the external validation to test the generalizability of our results, for example on DM2 and/or CHD patients without subthreshold depression, or outside the Dutch setting. Subsequently, studies are required to investigate the influence of the prediction model on decision making and patient outcomes. Consecutively, future research should evaluate whether the suggested enhanced vigilance strategies in combination with depression prevention programs that only target those with all four indicated predictors present and aim to reduce both anxiety and depressive symptoms, are cost-effective[65].

REFERENCES

1. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 2006;3:2011–30.

2. WHO. World Health Statistics 2017: Monitoring health for the SDGs. 2017.

3. Roy T, Lloyd CE. Epidemiology of depression and diabetes: A systematic review. J. Affect. Disord.; 2012;142:S8–21.

4. Rudisch B, Nemeroff CB. Epidemiology of comorbid coronary artery disease and depression. Biol. Psychiatry. 2003;54:227–40.

5. Lin EHB, Katon W, Von Korff M, Rutter C, Simon GE, Oliver M, et al. Relationship of depression and diabetes self-care, medication adherence, and preventive care. Diabetes Care. 2004;27:2154–60.

6. Gehi A, Haas D, Pipkin S. Depression and medication adherence in outpatients with coronary heart disease. Arch Intern Med. 2005;165:2508–13.

7. Ali S. The association between depression and health-related quality of life in people with type 2 diabetes: a systematic literature review. Diabetes. Metab. Res. Rev. 2010;26:75–89.

8. Rutledge T, Bittner V, Olson MB, Linke SE, Cornell CE, Eteiba W, et al. Depression and Cardiovascular Health Care Costs Among Women With Suspected Myocardial Ischemia (Women's Ischemia Syndrome Evaluation) Study. Jac. American College of Cardiology Foundation; 2009;53:176–83.

9. Bosmans JE, Adriaanse MC. Outpatient costs in pharmaceutically treated diabetes patients with and without a diagnosis of depression in a Dutch primary care setting. BMC Health Serv. Res. 2012;12:46.

10. Katon W, Lin EHB, Von Korff M, Ciechanowski P, Ludman E, Young B, et al. Integrating depression and chronic disease care among patients with diabetes and/or coronary heart disease: The design of the TEAMcare study. Contemp. Clin. Trials.; 2010;31:312–22.

11. Sullivan M, O'Connor P, Feeney P, Hire D, Simmons DL, Raisch D, et al. Depression Predicts All-Cause Mortality. Diabetes Care. 2012;35:1708–15.

12. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. Lancet; 2009;374:609–19.

13. National Collaborating Centre for Mental Health. Depression in adults with a chronic physical health problem. The NICE Guideline of Treatment and Management . 2010.

14. Chisholm D, Sanderson K, Ayuso-Mateos JL, Saxena S. Reducing the global burden of depression: Population-level analysis of intervention cost-effectiveness in 14 world regions. Br. J. Psychiatry. 2004;184:393–403.

15. Cuijpers P, van Straten A, Smit F, Mihalopoulos C, Beekman A. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. Am. J. Psychiatry. 2008;165:1272–80.

16. van Zoonen K, Buntrock C, Ebert DD, Smit F, Reynolds CF, Beekman ATF, et al. Preventing the onset of major depressive disorder: A meta-analytic review of psychological interventions. Int. J. Epidemiol. 2014;43:318–29.

17. Bower P, Gilbody S. Stepped care in psychological therapies: Access, effectiveness and efficiency. Narrative literature review. Br. J. Psychiatry. 2005;186:11–7.

18. van Dijk SEM, Pols AD, Adriaanse MC, Bosmans JE, Elders PJM, van Marwijk HWJ, et al. Cost-effectiveness of a stepped-care intervention to prevent major depression in patients with type 2 diabetes mellitus and/or coronary heart disease and subthreshold depression: design of a cluster-randomized controlled trial. BMC Psychiatry. 2013;13:128.

19. Davidson SK, Harris MG, Dowrick CF, Wachtler CA, Pirkis J, Gunn JM. Mental health interventions and future major depression among primary care patients with subthreshold depression. J. Affect. Disord.; 2015;177:65–73.

20. Pols AD, Van Dijk SE, Bosmans JE, Hoekstra T, Van Marwijk HWJ, Van Tulder MW, et al. Effectiveness of a stepped-care intervention to prevent major depression in patients with type 2 diabetes mellitus and/or coronary heart disease and subthreshold depression: A pragmatic cluster randomized controlled trial. PLoS One. 2017;12.

21. Bot M, Pouwer F, Ormel J, Slaets JPJ, de Jonge P. Predictors of incident major depression in diabetic outpatients with subthreshold depression. Diabet. Med. 2010;27:1295–301.

22. Pibernik-Okanovic M, Begic D, Peros K, Szabo S, Metelko Z. Psychosocial factors contributing to persistent depressive symptoms in type 2 diabetic patients: a Croatian survey from the European Depression in Diabetes Research Consortium. J. Diabetes Complications. 2008;22:246–53.

23. Badawi G, Pagé V, Smith KJ, Gariépy G, Malla A, Wang J, et al. Self-rated health: A predictor for the three year incidence of major depression in individuals with Type II diabetes. J. Affect. Disord. 2013;145:100–5.

24. Fisher L, Skaff MM, Mullan JT, Arean P, Glasgow R, Masharani U. A longitudinal study of affective and anxiety disorders, depressive affect and diabetes distress in adults with type 2 diabetes. Diabet. Med. 2008;25:1096–101.

25. Katon W, Ph JR, M EHBLMD, D SRHM, M PCMD, Evette J Ludman Ph, et al. Depression and Diabetes: Factors Associated With Major Depression at Five-Year Follow-Up. Psychosomatics. 2011;50:570–9.

26. Bot M, Pouwer F, Ormel J, Slaets JPJ, de Jonge P. Predictors of incident major depression in diabetic outpatients with subthreshold depression. Diabet. Med. England; 2010;27:1295–301.

27. Nefs G, Pouwer F, Denollet J, Pop V. The course of depressive symptoms in primary care patients with type 2 diabetes: Results from the Diabetes, Depression, Type D Personality Zuidoost-Brabant (DiaDDZoB) Study. Diabetologia. 2012;55:608–16.

28. Doyle F, McGee H, Delaney M, Motterlini N, Conroy R. Depressive vulnerabilities predict depression status and trajectories of depression over 1 year in persons with acute coronary syndrome. Gen. Hosp. Psychiatry.; 2011;33:224–31.

29. Spijkerman TA, Van Den Brink RHS, Jansen JHC, Crijns HJGM, Ormel J. Who is at risk of post-MI depressive symptoms? J. Psychosom. Res. 2005;58:425–32.

30. Pedersen SS, Denollet J, van Gestel YRBM, Serruys PW, van Domburg RT. Clustering of psychosocial risk factors enhances the risk of depressive symptoms 12-months post percutaneous coronary intervention. Eur. J. Cardiovasc.

Prev. Rehabil. 2008;15:203–9.

31. Ossola P, Paglia F, Pelosi A, De Panfilis C, Conte G, Tonna M, et al. Risk factors for incident depression in patients at first acute coronary syndrome. Psychiatry Res. 2015;228:448–53.

32. Kang H, Stewart R, Bae K, Kim S, Shin I, Hong Y, et al. Predictors of depressive disorder following acute coronary syndrome: Results from K-DEPACS and EsDEPACS. J. Affect. Disord.; 2015;181:1–8.

33. Kroenke K SR. The PHQ-9: a new depression diagnostic and severity measure. Psychiatr Ann. 2002;32:509–515.

34. Lamers F, Jonkers CCM, Bosma H, Penninx BWJH, Knottnerus JA, van Eijk JTM. Summed score of the Patient Health Questionnaire-9 was a reliable and valid method for depression screening in chronically ill elderly patients. J. Clin. Epidemiol. 2008;61:679–87.

35. Sheehan DV, Lecrubier Y SK. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatr. 1998;59:22–33.

36. Van Vliet IM, De Beurs E. Het Mini Internationaal Neuropsychiatrisch Interview (MINI): Een kort gestructureerd diagnostisch psychiatrisch interview voor DSM-IV-en ICD-10-stoornissen. Tijdschr. Psychiatr. 2007;49:393–7.

37. Meader N, Mitchell AJ, Chew-graham C, Goldberg D, Rizzo M, Bird V, et al. Case identification of depression in patients with chronic physical health problems: a diagnostic accuracy meta-analysis of 113 studies. 2011;808–20.

38. Moriarty AS, Gilbody S, McMillan D, Manea L. Screening and case finding for major depressive disorder using the Patient Health Questionnaire (PHQ-9): A meta-analysis. Gen. Hosp. Psychiatry; 2015;37:567–76.

39. van der Zwaan GL, van Dijk SEM, Adriaanse MC, van Marwijk HWJ, van Tulder MW, Pols AD, et al. Diagnostic accuracy of the Patient Health Questionnaire-9 for assessment of depression in type II diabetes mellitus and/or coronary heart disease in primary care. J. Affect. Disord. 2015;190:68–74.

40. Pols A, Schipper K, Overkamp D, van Marwijk H, van Tulder M, Adriaanse M. Patients' and practice nurses' perceptions of depression in patients with diabetes type 2 and/or coronary heart disease screened for

subthreshold depression. submitted.

41. Spinhoven P, Ormel J, Sloekers PP, Kempen GI, Speckens AE, Van Hemert AM. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. Psychol Med. 1997;27:363–70.

42. Robins LN, Helzer JE, Croughan J, Ratcliff KS. National Institute of Mental Health Diagnostic Interview Schedule. Arch Gen Psychiatry. 1981;38:381–9.

43. Kriegsman DMW, Penninx BWJH, Van Eijk JTM, Boeke a. JP, Deeg DJH. Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly. A study on the accuracy of patients' self-reports and on determinants of inaccuracy. J. Clin. Epidemiol. 1996;49:1407–17.

44. Twisk J. Different Methods to Analyse the Results of a Randomized Controlled Trial with More Than One Follow-Up Measurement. In: K. van Montfoort, J. Oud & WG, editor. Dev. Stat. Eval. Clin. Trials.; 2014. p. 177–93.

45. McCulloch CE NJ. Generalized Linear Mixed Models. Encyclopedia of Biostatistics. John Wiley & Sons; 2005.

46. Twisk JW. Applied longitudinal data analysis for epidemiology: a practical guide. Cambridge University Press; 2013.

47. Ali S, Stone MA, Peters JL, Davies MJ, Khunti K. The prevalence of co-morbid depression in adults with Type 2 diabetes: a systematic review and meta-analysis. Diabet. Med. 2006;23:1165–73.

48. Comijs HC, Nieuwesteeg J, Kok R, van Marwijk HW, van der Mast RC, Naarding P, et al. The two-year course of late-life depression; results from the Netherlands study of depression in older persons. BMC Psychiatry.; 2015;15:1.

49. de Boer MR, Waterlander WE, Kuijper L, Steenhuis I, Twisk J. Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate. Int. J. Behav. Nutr. Phys. Act. 2015;12:4.

50. Van Buuren, S; Groothuis-oudshoorn K. mice : Multivariate Imputation by Chained. J. Stat. Softw. 2011;45.

51. Rubin DB. Multiple imputation for nonresponse in surveys. John Wiley & Sons; 2004.

52. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. Eur. Urol. 2015;67:1142–51.

53. Heymans MW, Buuren S Van, Knol DL, Van W, Vet HCW De. Variable selection under multiple imputation using the bootstrap in a prognostic study. 2007;10:1–10.

54. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. BMC Med. Res. Methodol. 2009;9:57.

55. Reddy P, Philpot B, Ford D, Dunbar JA. Identification of depression in diabetes: the efficacy of PHQ-9 and HADS-D. Br. J. Gen. Pract. England; 2010;60:e239-45.

56. Cuijpers P, Koole SL, van Dijke A, Roca M, Li J, Reynolds CF. Psychotherapy for subclinical depression: meta-analysis. Br. J. psychiatry. 2014;205:268–74.

57. van't Veer-Tazelaar PJ, van Marwijk HWJ, van Oppen P, van Hout HPJ, van der Horst HE, Cuijpers P, et al. Stepped-care prevention of anxiety and depression in late life: a randomized controlled trial. Arch. Gen. Psychiatry. 2009;66:297–304.

58. Depressie N. M44 NHG-Standaard Depressie. Huisarts&Wetenschap. 2012;55:252–9.

59. Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: A meta-analysis. Cmaj. 2008;178:997–1003.

60. Moons K, Royston P, Vergouwe Y, Grobbee D, Altman D. Prognosis and prognostic research: what, why, and how? BMJ. 2009;338.

61. Kendler KS, Karkowski LM, Prescott CA. Stressful Life Events and Major Depression: Risk Period, Long-Term Contextual Threat, and Diagnostic Specificity. J. Nerv. Ment. Dis. 1998;186:661–9.

62. Assari S, Lankarani MM. Stressful life events and risk of depression 25 years later: Race and gender differences. Front. Public Heal. 2016;4:49.

63. Cuijpers P, Smit F, Willemse G. Predicting the onset of major depression in subjects with subthreshold depression in primary care: A prospective study. Acta Psychiatr. Scand. 2005;111:133–8.

64. Vink D, Aartsen MJ, Schoevers RA. Risk factors for anxiety and depression in the elderly: A review. J. Affect. Disord. 2008;106:29–44.

65. Steyerberg E, Moons KGM, van der Windt D, Hayden J, Perel P, Schroter S, et al. Prognosis research strategy PROGRESS series 3: prognostic model research. PLoS Med. 2013;10:e1001381.

Table 1 Patients' baseline characteristics at baseline in intervention group, care as usual group and total sample

Characteristics	Total sample (N=236)	Intervention (N=96)	Care as usual (N=140)
Female	107/236 (45.3)	42/96 (43.8)	65/140 (46.4)
Age, mean (SD)	67.5 (10.0)	67.8 (9.2)	67.3 (10.5)
Stressful life-event	112/210 (53.3)	48/89 (53.9)	64/121 (52.9)
Positive history of depression	113/210 (53.8)	54/89 (60.7)	59/121 (48.8)
ICPC diagnosis DM2 and/or CHD			
Diabetes Mellitus type 2 (DM2)	88/236 (37.3)	38/96 (39.6)	50/140 (35.7)
Coronary Heart Disease (CHD)	86/236 (36.4)	36/96 (37.5)	50/140 (35.7)
DM2 and CHD	62/236 (26.3)	22/96 (22.9)	40/140 (28.6)
More than 3 chronic diseases	98/210 (46.7)	38/89 (42.7)	60/121 (49.6)
PHQ-9 at baseline, mean (SD)	9.4 (3.2)	9.5 (3.1)	9.3 (3.2)
Anxiety HADS, mean (SD)	6.5 (3.8)	6.9 (3.7)	6.3 (3.9)
Depression HADS, mean (SD)	6.5 (3.8)	6.9 (3.9)	6.1 (3.7)
Marital status			
Married/living together	122/220 (55.5)	55 (61.1)	67/130 (51.5)
Single/divorced/widowed	98/220 (44.5)	35 (38.9)	63/130 (48.5)
Both parents born in the Netherlands	186/220 (84.5)	74/90 (82.2)	112/130 (86.2)
Rural residential area	99/236 (41.9)	42 (43.8)	57/140 (40.7)
Unemployed/sick	26/220 (11.8)	12/90 (13.3)	14/130 (10.8)
Level of education			
Low	89/220 (40.5)	33/90 (36.7)	56/130 (43.1)
Average	60/220 (27.3)	22/90 (24.4)	38/130 (29.2)
High	71/220 (32.3)	35/90 (38.9)	36/130 (27.7)
Current smoker	39/219 (17.8)	16/90 (17.8)	23/129 (17.8)
Alcohol use above norm	63/219 (28.8)	29/90 (32.2)	34/129(26.4)
Exercise under norm	141/219 (64.4)	56/90 (62.2)	85/129 (65.9)
BMI, mean (SD)	28.9 (6.1)	29.4 (6.8)	28.5 (5.6)
Locus of Control, mean (SD)	7.9 (4.2)	8.3 (4.2)	7.6 (4.1)
Social support, mean (SD)	36.3 (9.2)	35.8 (9.0)	36.7 (9.5)
Dysthymia	13/236 (5.5)	6/96 (6.3)	7/140 (5.0)
Onset of depression after age of 55	101/210 (48.1)	38/89 (42.7)	63/121 (52.1)

Numbers are percentages unless stated otherwise; Abbreviations: BMI = Body Mass Index; PHQ-9, Patient Health Questionnaire-9; HADS, Hospital Anxiety and Depression Scale; SD, Standard Deviation.

Table 2 Results of the mixed model and GEE long-term effectiveness analyses

Cumulative incidence of depression (n/N) %	Intervention	Care as usual	Corrected analyses*		Crude analyses	
Baseline	0	0	**OR (95%CI)**	**P-value**	**OR (95%CI)**	**P-value**
T6	(5/84) 6.0	(10/125) 8.0	0.82 (0.19; 3.51)	0.79	0.90 (0.32; 2.50)	0.84
T12	(9/82) 11.0	(12/118) 10.2	1.44 (0.46; 4.47)	0.53	1.20 (0.49; 2.92)	0.70
T24	(13/77) 16.9	(17/105) 16.2	1.23 (0.50; 3.02)	0.66	1.11 (0.51; 2.44)	0.79
Overall effect	n.a	n.a	1.37 (0.52;3.55)	0.52	1.11 (0.49;2.49)	0.80
PHQ mean (SD)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	9.53 (3.14)	9.28 (3.23)	**B (95%CI)**	**P-value**	**B (95%CI)**	**P-value**
T3	6.68 (4.55)	6.58 (4.21)	-0.39 (-1.52; 0.74)	0.50	-0.03 (-1.17; 1.11)	0.96
T6	6.10 (4.43)	6.12 (4.41)	-0.37 (-1.50; 0.76)	0.52	-0.17 (-1.30; 0.95)	0.76
T9	6.28 (4.31)	6.46 (4.51)	-0.48 (-1.62; 0.65)	0.40	-0.40 (-1.53; 0.73)	0.49
T12	6.60 (5.23)	6.29 (4.46)	-0.09 (-1.20; 1.02)	0.88	-0.03 (-1.13; 1.07)	0.96
T24	5.81 (4.76)	5.15 (4.33)	0.00 (-1.18; 1.19)	0.88	0.02 (-1.15; 1.19)	0.97
Overall effect	n.a	n.a	0.29 (-1.15; 0.58)	0.52	-0.13 (-0.99; 0.73)	0.77
Perceived recovery (%)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	n.a	n.a	**OR (95%CI)**	**P-value**	**OR (95%CI)**	**P-value**
T3	40.3%	49.5%	0.78 (0.42; 1.45)	0.44	0.64 (0.36; 1.15)	0.14
T6	48.8%	45.5%	1.46 (0.79; 2.69)	0.23	1.15 (0.65; 2.02)	0.64
T9	55.0%	48.7%	1.47 (0.79; 2.75)	0.22	1.30 (0.74; 2.30)	0.91
T12	55.6%	58.1%	1.04 (0.56; 1.92)	0.91	0.91 (0.51; 1.61)	0.74
T24	68.0%	57.1%	2.38 (1.21; 4.67)	0.01	2.04 (1.08; 3.87)	0.03
Overall effect	n.a	n.a	1.32 (0.87; 2.00)	0.19	1.10 (0.75; 1.62)	0.61
HADS-A mean (SD)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	6.91 (3.74)	6.25 (3.90)	**B (95%CI)**	**P-value**	**B (95%CI)**	**P-value**
T3	6.35 (4.04)	6.29 (3.97)	-0.27 (-1.13; 0.60)	0.54	-0.13 (-1.00; 0.74)	0.76
T6	5.70 (4.10)	6.63 (4.00)	-1.04 (-1.91; -0.18)	0.02	-1.04 (-1.91; -0.18)	0.02
T9	6.16 (4.24)	6.03 (4.04)	-0.49 (-1.35; 0.38)	0.27	-0.45 (-1.31; 0.42)	0.31
T12	5.77 (4.69)	5.83 (3.99)	-0.50 (-1.37; 0.38)	0.27	-0.43 (-1.31; 0.44)	0.33
T24	5.45 (4.46)	5.06 (3.90)	-0.59 (-1.50; 0.31)	0.20	-0.48 (-1.38; 0.43)	0.30
Overall effect	n.a	n.a	-0.59 (-1.23; 0.06)	0.08	-0.52 (-1.17; 0.13)	0.12
HADS-D mean (SD)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	6.93 (3.87)	6.11 (3.73)	**B (95%CI)**	**P-value**	**B (95%CI)**	**P-value**
T3	6.14 (4.16)	6.21 (3.87)	-0.26 (-1.12; 0.60)	0.55	-0.29 (-1.15; 0.56)	0.51
T6	5.82 (3.79)	5.75 (4.03)	-0.22 (-1.07; 0.64)	0.62	-0.32 (-1.18; 0.53)	0.46
T9	6.36 (4.04)	6.07 (4.08)	-0.21 (-1.06; 0.65)	0.63	-0.24 (-1.09; 0.61)	0.58
T12	6.09 (4.20)	6.11 (4.22)	-0.41 (-1.27; 0.46)	0.36	-0.50 (-1.36; 0.36)	0.26
T24	5.59 (4.66)	4.92 (3.90)	-0.41 (-1.30; 0.48)	0.37	-0.48 (-1.37; 0.41)	0.29
Overall effect	n.a	n.a	-0.30 (-0.94; 0.33)	0.35	-0.37 (-1.00; 0.26)	0.25

Abbreviations: 95%CI, 95% Confidence Interval; HADS-A, Hospital Anxiety and Depression Scale-Anxiety; HADS-D, Hospital Anxiety and Depression Scale-Depression; n.a, not applicable; PHQ-9, Patient Health Questionnaire-9;

*Corrected for: baseline values of the outcome, age, gender, marital status, employment status, level of education, co-existence of DM2 and CHD, alcohol use, number of depressive episodes in history and age of onset of depression. The baseline value of the outcome is not added as an extra variable in the corrected analyses of the overall effects since it is already incorporated in the crude overall analyses.

Table 3 Multivariable prediction model of incident depression during two-year follow-up

Predictor	RC	OR	95% CI	P-value
Female sex		-	-	-
Age		-	-	-
Somatic disorder DM2 CHD DM2and CHD		-	-	-
History of depression		-	-	-
Baseline depression scores	0.32 p.p.i.	1.37	1.20; 1.55	0.00
Baseline anxiety scores	0.12 p.p.i.	1.13	1.02; 1.25	0.01
Stressful life-event in past year	0.74	2.10	1.02; 4.32	0.04
>3 chronic illnesses	0.78	2.19	1.12; 4.25	0.02
Randomization status I vs C	0.14	1.15	0.58; 2.29	0.68

RC regression coefficient; p.p.i. per point increase; 95% CI 95% confidence interval; OR odds ratio, an OR > 1 reflects a higher probability the outcome an incident depression and an OR < 1 reflects a lower probability compared with the reference category. OR estimated after multiple imputation (n = 25 datasets) with p-value of 0.157. Linear predictor corrected after bootstrapping = -4.1147 + 0.131* Randomization status + 0.7167* >3 chronic illnesses + 0.680* stressful life-event in past year + 0.1118* baseline anxiety scores + 0.2868* baseline depression scores

SUPPLEMENTARY INFORMATION

Contributors

AP constructed the design of this study, performed all statistical analyses and drafted the manuscript. MCA, MvT, HvM constructed the design of the study and revised the manuscript. JB and SvD constructed the design of the Step-Dep study and revised the manuscript. MH collaborated on the statistical analyses and revised the manuscript. The final manuscript was read and approved by all authors.

Funding

This study is funded by ZonMw, the Netherlands Organisation for Health Research and Development (project number 80-82310-97-12110). The sponsor had no role in the design and conduct of the present study or in the writing of the manuscript.

Competing Interests

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Trial registration and ethical approval

The study was performed in accordance with the declaration of Helsinki (2008) and the Dutch Medical Research involving Human Subjects Act (WMO). The protocol was approved by the medical ethics committee of the VU University Medical Centre (NL39261.029.12, registration number 2012/223), and registered in the Dutch Trial Register (registration number 3715).

Data sharing

Full dataset and statistical code is available from the corresponding author. Consent was not obtained but the presented data are anonymised and risk of identification is low.

Transparency

The corresponding author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; no important aspects of this study have been omitted; and any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Acknowledgements

The authors would like to thank Marcella van der Linden, Lucca Vledder and Mieke Schlattmann for their contribution in the data collection for this study and Jos Twisk for his help in the long-term effectiveness analyses. We also would like to thank all the participating general practices and the research networks of general practitioners (ANH, THOON and LEON) for their participation and collaboration in the implementation and execution of the study. Furthermore, this study has been possible thanks to all Step-Dep participants.

SUPPLEMENTARY FILES

S1 Original protocol

S2 TRIPOD statement checklist

TRIPOD

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	4-5
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	5
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	6
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	6
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	6
	5b	D;V	Describe eligibility criteria for participants.	6
	5c	D;V	Give details of treatments received, if relevant.	7
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	6
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	6
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	7
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	6
Sample size	8	D;V	Explain how the study size was arrived at.	Ref protocol
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	8
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	7-8
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	8-9
	10c	V	For validation, describe how the predictions were calculated.	n.a.
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	9
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	n.a.
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	n.a.
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	n.a.
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	10
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	10, table 1
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	n.a.
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	n.a.
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	n.a.
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	Table 3
	15b	D	Explain how to the use the prediction model.	10-11
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	10-11
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	n.a.
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	12
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	n.a.
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	12-13
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	13-14
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	24-26
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	24

TRIPOD Checklist: Prediction Model Development and Validation

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

BMJ Open

Two-year effectiveness of a stepped-care depression prevention intervention and predictors of incident depression in primary care patients with diabetes type 2 and/or coronary heart disease and subthreshold depression; data from the Step-Dep cluster randomized controlled trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-020412.R3
Article Type:	Research
Date Submitted by the Author:	28-Jun-2018
Complete List of Authors:	Pols, Alide; VU University, Department of Health Sciences Adriaanse, Marcel; Institute of Health Sciences, Vrije Universiteit Amsterdam Tulder, Maurits; University of Amsterdam, Health Sciences Heymans, Martijn; VU University Medical Center, Department of Epidemiology and Biostatistics; VU University, Department of Health Sciences Section Methodology and Applied Biostatistics Faculty of Earth and Life Sciences Bosmans, J; VU University Amsterdam, Department of Health Sciences van Dijk, Susan; Vrije Universiteit Amsterdam, Department of Health Sciences Van Marwijk, H; University of Brighton, Division of Primary Care and Public Health, Brighton and Sussex Medical School, Mayfield House
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Cardiovascular medicine, Diabetes and endocrinology, Mental health, General practice / Family practice
Keywords:	Depression & mood disorders < PSYCHIATRY, DIABETES & ENDOCRINOLOGY, Coronary heart disease < CARDIOLOGY, PREVENTIVE MEDICINE

Two-year effectiveness of a stepped-care depression prevention intervention and predictors of incident depression in primary care patients with diabetes type 2 and/or coronary heart disease and subthreshold depression; data from the Step-Dep cluster randomized controlled trial

Alide D Pols*[1,2], Marcel C Adriaanse[1], Maurits W van Tulder[1], Martijn W Heymans[3], Judith E Bosmans[1], Susan E van Dijk[1], Harm W J van Marwijk[2,4]

[1] Department of Health Sciences, Amsterdam Public Health research institute, Faculty of Science, Vrije Universiteit, Amsterdam, The Netherlands

[2] Department of General Practice & Elderly Care Medicine, Amsterdam Public Health research institute, VU University Medical Centre, Amsterdam, The Netherlands

[3] Department of Clinical Epidemiology and Biostatistics VU Medical Centre Amsterdam, The Netherlands

[4] Division of Primary Care and Public Health, Brighton and Sussex Medical School, Mayfield House, University of Brighton, Brighton, United Kingdom

* Corresponding author: a.d.pols@vu.nl

Keywords: major depressive disorder, subthreshold depression, diabetes mellitus type 2, coronary heart disease, effectiveness, stepped care, prediction model

ABSTRACT

Introduction Major depressive disorders (MDD), diabetes mellitus type 2 (DM2) and coronary heart disease (CHD) are leading contributors to the global burden of disease and often co-occur.

Objectives To evaluate the two-year effectiveness of a stepped-care intervention to prevent MDD compared to usual care and to develop a prediction model for incident depression in DM2 and/or CHD patients with subthreshold depression.

Methods Data of 236 Dutch primary care DM2/CHD patients with subthreshold depression (Patient Health Questionnaire 9 (PHQ-9) score ≥6, no current MDD according to the Mini International Neuropsychiatric Interview (DSM-IV criteria)), who participated in the Step-Dep trial were used. A PHQ-9 score of ≥10 at minimally one measurement during follow-up (at 3, 6, 9, 12 and 24 months) was used to determine the cumulative incidence of MDD. Potential demographic and psychological predictors were measured at baseline via web-based self-reported questionnaires and evaluated using a multivariable logistic regression model. Model performance was assessed with the Hosmer–Lemeshow test, Nagelkerke's R^2 explained variance and Area Under the Receiver Operating Characteristic curve (AUC). Bootstrapping techniques were used to internally validate our model.

Results 192 patients (81%) were available at two-year follow-up. The cumulative incidence of MDD was 97/192 (51%). There was no statistically significant overall treatment effect over 24 months of the intervention (OR 1.37; 95% CI 0.52; 3.55). Baseline levels of anxiety, depression, the presence of >3 chronic diseases and stressful life-events predicted the incidence of MDD (AUC 0.80 interquartile range (IQR) 0.79-0.80; Nagelkerke's R^2 0.34 IQR 0.33-0.36).

Conclusion A model with four factors predicted depression incidence during two-year follow-up in patients with DM2/CHD accurately, based on the AUC. The Step-Dep intervention did not influence the incidence of MDD. Future depression prevention programs should target patients with these four predictors present, and aim to reduce both anxiety and depressive symptoms.

TRIAL REGISTRATION NUMBER

Dutch Trial Register NTR3715 http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=3715

STRENGTHS AND LIMITATIONS OF THIS STUDY

- This study provides a prediction model of incident MDD in DM2 and/or CHD patients with subthreshold depression, which could assist healthcare providers in its detection and facilitate targeting indicated prevention to highest risk patients
- Only predictors that are readily available or easily obtained in practice were used in the multivariable model, which enhances the practical use of the model
- This study had a relatively long follow-up and outcomes were frequently measured, whereas drop-out rates were relatively low and missing values imputed
- The relatively small study population might have caused over-optimism of the prediction model, but an internal validation procedure with bootstrapping techniques showed that this risk was minor
- Data were derived from a RCT, but statistically non-significant intervention effects for incident MDD at both 12- and 24-months follow-up justify using the Step-Dep population as a cohort

INTRODUCTION

Depression is a major and increasing contributor to the global burden of disease[1], whereas coronary heart disease (CHD) and diabetes mellitus type 2 (DM2) rank among the leading causes of morbidity and mortality worldwide[2]. Comorbid depression in patients with DM2 and/or CHD is common[3,4] and has detrimental effects on self-care and medication adherence[5,6], quality of life[7], health status and increases healthcare costs[8,9] and mortality[10,11]. Despite its negative impact, many cases of depression go unrecognized in primary care[12], especially in patients with chronic diseases like DM2 and/or CHD[13]. Additionally, about one-third of those recognized and treated does not respond to current approaches, and over half of those who experience a first onset of a major depressive episode will experience one or more recurrences[14].

Given the significant burden of disease of depression, its poor recognition and the limited effect of current treatment options for it, it would be of great value if incident cases could be averted by early detection and preventive treatment of patients at risk ('indicated prevention'). Meta-analyses have shown that preventive psychological interventions can overall reduce the incidence of MDD in comparison to control groups[15,16]. Offering preventive psychological interventions in a stepped-care format could be an efficient approach, as patients start with minimally intensive evidence-based treatments and only those who do not improve adequately, step up to a treatment of higher intensity[17]. Recently, we conducted a randomized controlled trial in which we evaluated whether a pragmatic nurse-led stepped-care program was effective in reducing the incidence of MDD at 12-months of follow-up in comparison with usual care among patients with DM2 and/or CHD and subthreshold depression (Step-Dep)[18]. Subthreshold depression entails clinically relevant depressive symptoms without fulfilling the criteria for MDD and is a known important risk factor for depression[15,19]. We demonstrated that the Step-Dep intervention was not superior to usual care and the overall cumulative incidence of MDD was lower than expected after one year [20]. However, it may be possible that the follow-up period was too restricted to demonstrate the potential health benefits of the stepped-care program over usual care, or the presence of subthreshold depression alone posed a lower than expected prior risk of MDD in our DM2 and/or CHD population.

Identifying additional major risk factors of incident depression in patients with DM2 and/or CHD might facilitate targeting indicated prevention to patients with highest risk, but also potentially aid in its detection. In patients with DM2, several longitudinal studies have been conducted to determine risk factors for comorbid incident depression. However, these studies have rendered heterogeneous results, due to small patient samples (<80 at follow-up)[21,22], analyses of single factors only[23,24], the use of mixed samples of type 1 diabetes and DM2[25], patients with either no MDD at baseline[23,26] or both with and without depression at baseline[22,24,25,27], and differences across community[23,24], primary care[25,27] and secondary care settings[22,26]. In patients with CHD, the only available longitudinal data are derived from studies in patients with acute coronary syndrome followed-up after hospital discharge[28–32]. Predictors that were repeatedly identified in DM2 or CHD studies were: depression severity at baseline[21,22,25,28,31,32], history of depression[25,27,29], female sex[24,27,29,31] and baseline anxiety levels[21,30,31]. However, data of patients with both DM2 and CHD, non-acute CHD or within primary care settings are scarce. The goal of the present study was twofold: (1) to evaluate the two-year effectiveness of a nurse-led stepped-care intervention to prevent MDD as compared to usual care (Step-Dep); and to (2) develop a prediction model for incident depression during two-year follow-up in primary care patients with DM2 and/or CHD and subthreshold depression.

METHODS

Design

Data of the Step-Dep cluster randomized controlled trial were used. Step-Dep was conducted in 27 general practitioner (GP) practices in three regions in the Netherlands (Amsterdam, Leiden, Twente), between January 2013 and November 2016, including recruitment and two years of follow-up. A statistician blinded to the characteristics of the GP practices performed the randomization of GP practices using a computer generated list of random numbers. Randomization was done at the level of the GP practice, which corresponds to the participating practice nurse, to avoid contamination between the treatment groups, and was stratified for size (less or more than 5000 patients). The study was performed in accordance with the declaration of Helsinki (2008) and the Dutch Medical Research involving Human Subjects Act (WMO). The protocol was approved by the medical ethics committee of the VU University Medical Centre (NL39261.029.12, registration number 2012/223), and registered in the Dutch Trial Register (NTR3715 http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=3715) The outcomes of the two-year effectiveness of the Step-Dep study and predictors of incident depression were not pre-specified in designing the study. Further details on the methods and design of the Step-Dep study have been published elsewhere[18].

Patient and Public Involvement

Patients were not involved in determining the design, the recruitment to or conduct of the study. The medical ethics committee of the VU University Medical Centre assessed the burden of the intervention and participation in the study in general as acceptable for patients. The burden of and satisfaction with the intervention were assessed in a process evaluation with 15 patients. All patients are thanked in the acknowledgements section. Results of the study will be disseminated by letter to all participants.

Patients

Included patients were aged 18 years or more who had an International Classification of Primary Care (ICPC) diagnosis of DM2 and/or CHD and had subthreshold depression identified by screening. Patients with a Patient

Health Questionnaire 9 (PHQ-9; range 0-27 with higher scores indicating more severe depressive symptoms) score of six or higher[33,34], and no major depressive disorder according to the Mini International Neuropsychiatric Interview (MINI)[35,36], were considered to have subthreshold depression. Exclusion criteria were cognitive impairment, psychotic illnesses, a terminal illness, the use of anti-depressant medication, a history of suicide attempt(s), loss of significant other in the past six months, visual impairment, current pregnancy, bipolar disorder, borderline personality disorder or any difficulties completing written questionnaires or visiting the primary care center. A total of 236 patients gave informed consent to participate.

Outcome measure

The outcome measure used was an incident depression (yes/no) defined as a PHQ-9 score of ≥10 at minimally one moment during follow-up (measured at 3, 6, 9, 12 and 24 months after baseline). The PHQ-9 is a widely used and validated instrument that performs well in patients with chronic medical illnesses both as dichotomous diagnosis of major and minor depression and a continuous severity score[34,37]. A cut-off of ≥10 has been shown to be the optimum cutoff for major depression[38], also in this patient group [39]. PHQ-9 was self-reported with web-based or written questionnaires. When these web-based or written questionnaires were not completed, the PHQ-9 was administered by telephone by trained research-assistants, blinded to randomization status.

Potential predictors

The selection of the potential predictors was based on a thorough literature search. Predictors of incident depression that were identified in multiple studies in patients with DM2 or CHD and are routinely available or easily obtained in daily GP practice were used. Additionally, we chose the presence of multiple chronic diseases[24] and stressful life-events[28] although they were identified in single studies only, as these were also indicated as causes of depression by patients and practice nurses in semi-structured interviews as part of the process evaluation of Step-Dep[40], and age[23].

Apart from GP information system derived data on *sex, age* and *ICPC diagnosis of DM2 and/or CHD*, demographics and psychological factors were measured at baseline via web based (or written if preferred) self-reported

questionnaires. To take possible effects of the intervention into account, we included randomization status in the selection models as well. Patients in the intervention arm were offered a stepped care prevention program, and patients in the control arm received care as usual during one year. The stepped care intervention consisted of four sequential but flexible treatment steps, each lasting three months; 1) watchful waiting, 2) guided self-help, 3) problem solving treatment and 4) referral to a general practitioner. After each step, patients with a persisting PHQ-9 score of six or more were offered the next treatment step of the intervention. *Baseline depression levels* were measured with the PHQ-9[33,34]. *Baseline anxiety levels* were measured with the Hospital Anxiety and Depression Scale Anxiety (HADS-A; range 0-21 with higher scores indicating more severe anxiety)[41]. *History of depression* and s*tressful life-events were* self-reported using a subset of the Diagnostic Interview Schedule (DIS)[42]. *Number of co-morbid chronic illnesses* was measured using the self-reported Dutch Questionnaire Chronic Illnesses[43]. This was dichotomized using the median in our sample: three or less versus more than three chronic diseases.

Statistical analyses

The two-year effectiveness of the intervention on the primary and secondary outcomes was analyzed according to the intention to treat principle. Generalised Estimating Equations (GEE) were used for binary outcome variables, and linear mixed models for longitudinal data were used for continuous outcome variables[44]. For each outcome an overall effect over time and separate effects at different time points were estimated by taking time into account as a categorical variable (with five categories: 0-3 months, 3-6 months, 6-9 months, 9-12 months and 12-24 months of follow-up)[45,46]. The main analyses consisted of fully corrected models that were corrected for baseline values of the respective outcome and additionally included the covariates gender[47], age[48], and any other possible confounding variable on which the treatment groups differed at baseline (marital status, employment status, level of education, co-existence of DM2 and CHD, alcohol use, number of depressive episodes in history and age of onset of depression), based on absolute baseline differences judged by the researchers rather than statistical testing[49]. For these analyses, STATA version 14 was used.

Missing data were imputed using multiple imputation according to the Multivariate Imputation by Chained Equations (MICE) algorithm[50] in SPSS version 23. For the imputations, missing at random (MAR) was assumed.

Variables that were associated with missing data and variables that were associated with the outcome, were identified and included in the imputation model. Also, all variables in the analysis model (potential predictors and outcome) were included. The number of imputed datasets was 25 based on the proportion of cases with incomplete measurements; 24%. The subsequent analyses were performed on pooled data according to Rubin's rules[51].

Prediction model

We created a multivariable logistic regression model in SPSS 23 from the baseline variables estimating the probability of having at least one major depression (PHQ ≥10) during the two-year assessment. To calculate the number of potential predictors for developing the prediction model, we used the criterion of 10 events per variable. Continuous variables were checked for linearity with the outcome using spline regression curves and linearity was confirmed. All variables were entered into the logistic model and tested for statistical significance in the presence of the total set of predictors. Individually, the least significant predictor (P-value>0.157, as recommended in the TRIPOD statement. [52] Wald statistic was removed, and the model was refit (backward selection). Randomization status was maintained in the model. This was repeated until we reached a statistical model that only included statistically significant predictors. This was repeated with p-values of 0.05. We also compared the results with complete case analysis (CCA), i.e., all patients with missing data were excluded from the analyses.

We checked the performance of the model with regard to the goodness of fit (Hosmer–Lemeshow test), the explained variation and the discriminative ability of the model. The Nagelkerke's R^2 explained variation is the extent to which the outcome can be predicted by the predictors in the model in current datasets. The discriminative ability is reflected by the area under the receiver operating characteristics curve (AUC). Bootstrapping techniques were used to internally validate our model, i.e., to simulate the performance with respect to the explained variance and the AUC in comparable patient datasets[53]. After that, we calculated the linear predictor of the bootstrapped model with an adjusted intercept and regression coefficients corrected for the

shrinkage factor. Performance measures were assessed in each imputed dataset and results were summarized using median values [54]. All analyses were done with SPSS version 23.0 and R software.

RESULTS

Participants

The baseline characteristics of the study population are presented in Table 1. Of the 236 patients included in Step-Dep, 192 patients (81%) completed two years of follow-up. A flowchart of participants through the first 12 months of the Step-Dep study has been published elsewhere[20]. At 24 months of follow up 18 additional patients dropped out (two for unknown motives, seven due to time considerations, four were deceased, three too frail, two unable to contact). We compared the baseline characteristics of patients with missing data to those without. Patients with missing data were more often living alone (61% vs 41%), but no other differences between these groups were found.

There was no statistically significant overall treatment effect over 24 months of the intervention (OR 1.37; 95% CI 0.52 to 3.55), nor at any of the time-points. There were no significant differences in PHQ-9 scores between the study groups at any time-point and the course of PHQ-9 scores over time did not differ significantly between the groups. Results are shown in Table 2. The statistically non-significant intervention effects for incident MDD at both 12-months[20] and 24-months of follow-up justify using the Step-Dep population as a cohort.

Prediction model

The cumulative incidence during two-year follow-up was 97/192 (51%). The multivariable models using p=0.05 and p=0.157[52] were identical. The final model consisted of four predictors: level of anxiety, level of depression, presence of more than three chronic diseases and having suffered a stressful life-event in the past year. This model performed well (Hosmer–Lemeshow test p=0.12 and median of pooled Nagelkerke's R^2 explained variance 0.34 interquartile range (IQR) 0.33-0.36) with good discriminative properties (median of the pooled AUC 0.80 IQR 0.79-0.80). In a CCA with p=0.05, the same predictors remained. In a CCA using p=0.157 [52], the categorical variable DM2/CHD/both also remained.

The risk of an incident MDD during two years of follow-up more than doubled when either more than three chronic diseases were present or a patient had suffered a stressful life-event in the past year. Both higher

depression and anxiety levels at baseline increased the risk of MDD with each incremental point on the PHQ-9 of HADS scales respectively. One point higher on the PHQ-9 at baseline, resulted in a 1.37 higher risk of developing MDD during two years, compared to 1.13 for increasing anxiety levels. With regard to the internal validation of the model, the calibration slope (or shrinkage factor to correct regression coefficients of the original model) was 0.92 IQR 0.91-0.92, the median explained variance was 31% IQR 0.29-0.32 and the AUC 0.78 IQR 0.77-0.78. This means that after corrections for over-optimism, both the performance and discriminative properties of the model remained good. Results are shown in Table 3.

DISCUSSION

This study showed that the Step-Dep intervention was not more effective than usual care in the prevention of MDD at two years of follow-up. The risk of incident MDD during two years of follow-up among patients with DM2 and/or CHD and subthreshold depression, was increased by higher baseline levels of anxiety and depression, the presence of more than three chronic diseases and having suffered a stressful life-event in the past year. This risk was not influenced by a stepped-care intervention aimed at preventing MDD.

Our findings have to be viewed in the context of strengths and limitations of this study. Strengths are its relatively long follow-up with frequent outcome measurements and low drop-out rates. In addition, missing values were imputed using multiple imputation techniques. We only used predictors that are readily available or easily obtained in practice, which enhances the practical use of the model in primary care consultations. Furthermore, testing a multivariable model instead of single factors appointed only the most relevant predictors, which rendered a simple model that is manageable in its use. There were limitations to this study. First, the study population was relatively small, which might have caused over-optimism of the prediction model. This means that it predicts the outcome better in the sample used to develop the model than in new samples, potentially restricting its external validity. However, an internal validation procedure with bootstrapping techniques showed that this risk was minor. Second, we used data derived from a RCT instead of a cohort, which potentially limits the generalizability of our results. Third, we evaluated a limited number of predictors in this study and genetic and other biological risk indicators, for example, were not included. This was due to the relatively small population size and our pre-selection criteria for potential predictors: predictors had to be both identified before in multiple studies and easily obtainable in GP practice . Finally, in this study, the use of the PHQ-9 with a cut-off score of 10 or more rendered a higher cumulative incidence of depression than the MINI. This could be explained by the fact that the PHQ-9 was measured more frequently than the MINI. Also, the PHQ-9 was self-reported instead of administered with a diagnostic interview by a trained research assistant. However, it is possible that depression was sometimes over-diagnosed with the PHQ-9 due to potential overlap of (somatic) symptoms of the chronic disease and those of depression[55].

In a previous publication we have hypothesized the causes for the lack of effect of the Step-Dep intervention as compared to care as usual in preventing incident MDD at 12 months of follow-up[20], which we assume also explain the lack of effect at 24 months of follow-up. In summary, a first explanation could be that subthreshold depression was potentially over-diagnosed in our population, whereas stepped-care may be more effective in patients with more severe symptoms[56]. Secondly, fewer patients than expected were treated with the more intensive treatment steps. This was partly caused by the fact that a considerable proportion of patients did not want to start one or more of the treatment steps. This may indicate that our program did not sufficiently match their need for care. Furthermore, this was in part due to the low PHQ-9 scores of 6.7 on average at three months after baseline measurements, which made only a relatively small proportion of the patients eligible for more intensive treatment steps. The drop in PHQ-9 scores between baseline and three months of follow-up in both groups exceeded the expectations of spontaneous recovery alone[57]. It is unlikely that either of the groups received any specific treatment during this period. The Step-Dep program entailed an initial period of watchful waiting and Dutch primary care clinical guidelines recommend a similar waiting period before starting treatment for subthreshold depression[58]. Additionally, screening for depression alone does not change the management of depression in primary care[59]. We argue that the decrease in depressive symptoms may partly be caused by attention, regression to the mean, or patients' self-insight into their mental symptoms and problems. Finally, depressive and anxiety symptoms slightly improved over time in both groups, possibly indicating that usual care is already of reasonable quality and, therefore, the room for improvement for new interventions over usual care may be limited.

We observed a remarkable drop between baseline and three months in the PHQ-9, but not for the HADS-D. We can only speculate about this difference in drop between PHQ9 and HADS-D at three months. Currently we have no solid explanation for this difference. There is a possibility of a statistical artefact. The PHQ9 is made to align with DSM diagnostic symptoms of depression irrespective of the co-morbid presence of physical conditions while the HADS-D should be robust for physical illnesses and perhaps measures a broader construct (for instance, 'I can laugh and see the funny side of things'). We do think that the different sensitivity of these instruments have minimal implications, if at all, fort the intervention algorithm of the Step care approach. In the StepDep effectiveness study [20] we used the MINI, the PHQ9, the HADS-D and HADS-A to look at the differences in

incident major depression and depression and anxiety levels respectively. All instruments used are valid and reliable. We found no statistically significant differences at any time point nor a statistically significant difference in the course of incident MDD or depression and anxiety symptom levels over time between the groups. In other words, the slope of the different outcomes over time were virtually the same.

Our multivariable model consisted of four predictors of MDD incidence. Firstly, baseline depression severity level is the most frequently found and often strongest predictor of incident depression in other studies in patients with DM2[21,22,25] or CHD[28,31,32]. In line with these findings, in our model a clinically relevant baseline difference in depressive symptoms of five points on the PHQ-9, translated to an almost five times increased risk of developing a MDD during two years. This factor was used as a continuous variable in which the severity level predicts the occurrence of a depressive episode, which supports the concept of a gradual risk of depression. Secondly, the anxiety level at baseline was an important predictor of MDD. Anxiety has been frequently appointed as an important risk factor for depression in DM2[21] and CHD populations[30,31]. Predictors are not necessarily etiological factors[60]. Nonetheless, as anxiety is also known for its high comorbidity with depression, the assumption that reducing anxiety will have a positive effect on depressive symptoms and MDD incidence seems defendable. Thirdly, the risk the occurrence of stressful life-events pose, has been demonstrated before in patients with CHD[28]. Although most of our knowledge on the role of stressful life-events as predictors of depression cover a short period of time[61], more recent research has shown their long-term risk[62]. This would imply that healthcare providers should not only be temporarily alert on the negative influence on mental health of stressful life-events, but should also be aware of deferred effects. Fourthly, the presence of more than three chronic diseases was identified as a predictor of MDD in our study, in concordance with results in a DM2 population of Fisher et al.[24] Interestingly, the presence of either DM2, CHD or both was not a predictor in our study, which suggests that these patients are at the same risk of incident depression. As all included patients in Fisher's and our study had at least one chronic disease, a discrimination between the predictive values of no chronic disease versus only one versus multiple chronic diseases could not be made. The specific importance of an increased number of diseases as opposed to the risk of a chronic disease has also been demonstrated previously in a primary care population with subthreshold depression[63] and several elderly populations[64]. Why the number of diseases would matter in itself, can perhaps be understood from findings from qualitative interviews. Step-Dep patients

explained that chronic diseases indirectly lead to depression, as they diminish future perspectives and cause disability[40], which might be subjective to a certain "threshold" burden of disease. Finally, in contrast to findings in multiple other studies, female sex[24,27,29,31]and a history of depression[25,27,29] did not predict incident MDD in our study. These factors were also not univariately associated with incident depression in our data. A history of depression was self-reported in our study. Perhaps patients over-reported this, as it was not required that they received treatment for this depressive episode, which might explain the lack of an univariate correlation with incident depression.

The model rendered in this study had good discriminative properties with an AUC of 0.80 with the use of only four predictors that are relatively easily obtained by the GP. This makes this prediction model practically viable. It could assist as a tool to both improve the (early) recognition of depression in primary care patients with DM2 and/or CHD and indicate which patients need further care. As chronic care in the Netherlands is being delegated more and more to primary care practice nurses, such a tool might prove useful in their and the GPs' regular check-ups. In practice, this would not only entail that in patients with DM2 and/or CHD, GPs and practice nurses standardly inquire about symptoms of depression and anxiety during regular checkups, but also that in those with multiple chronic diseases next to their DM2 or CHD, who suffered a recent stressful life-event, the presence and course of depressive and anxiety symptoms are assessed and monitored over time with, for example, the PHQ-9 and HADS. Whenever depressive or anxiety symptoms are clinically severely elevated or significantly deteriorate over time, treatment should be offered according to the patients' need for care. By reducing both depressive and anxiety symptoms, perhaps MDD and its negative consequences can be averted.

Future research should focus on the external validation to test the generalizability of our results, for example on DM2 and/or CHD patients without subthreshold depression, or outside the Dutch setting. Subsequently, studies are required to investigate the influence of the prediction model on decision making and patient outcomes. Consecutively, future research should evaluate whether the suggested enhanced vigilance strategies in combination with depression prevention programs that only target those with all four indicated predictors present and aim to reduce both anxiety and depressive symptoms, are cost-effective[65].

REFERENCES

1. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 2006;3:2011–30.

2. WHO. World Health Statistics 2017: Monitoring health for the SDGs. 2017.

3. Roy T, Lloyd CE. Epidemiology of depression and diabetes: A systematic review. J. Affect. Disord.; 2012;142:S8–21.

4. Rudisch B, Nemeroff CB. Epidemiology of comorbid coronary artery disease and depression. Biol. Psychiatry. 2003;54:227–40.

5. Lin EHB, Katon W, Von Korff M, Rutter C, Simon GE, Oliver M, et al. Relationship of depression and diabetes self-care, medication adherence, and preventive care. Diabetes Care. 2004;27:2154–60.

6. Gehi A, Haas D, Pipkin S. Depression and medication adherence in outpatients with coronary heart disease. Arch Intern Med. 2005;165:2508–13.

7. Ali S. The association between depression and health-related quality of life in people with type 2 diabetes: a systematic literature review. Diabetes. Metab. Res. Rev. 2010;26:75–89.

8. Rutledge T, Bittner V, Olson MB, Linke SE, Cornell CE, Eteiba W, et al. Depression and Cardiovascular Health Care Costs Among Women With Suspected Myocardial Ischemia (Women's Ischemia Syndrome Evaluation) Study. Jac. American College of Cardiology Foundation; 2009;53:176–83.

9. Bosmans JE, Adriaanse MC. Outpatient costs in pharmaceutically treated diabetes patients with and without a diagnosis of depression in a Dutch primary care setting. BMC Health Serv. Res. 2012;12:46.

10. Katon W, Lin EHB, Von Korff M, Ciechanowski P, Ludman E, Young B, et al. Integrating depression and chronic disease care among patients with diabetes and/or coronary heart disease: The design of the TEAMcare study. Contemp. Clin. Trials.; 2010;31:312–22.

11. Sullivan M, O'Connor P, Feeney P, Hire D, Simmons DL, Raisch D, et al. Depression Predicts All-Cause Mortality. Diabetes Care. 2012;35:1708–15.

12. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. Lancet; 2009;374:609–19.

13. National Collaborating Centre for Mental Health. Depression in adults with a chronic physical health problem. The NICE Guideline of Treatment and Management . 2010.

14. Chisholm D, Sanderson K, Ayuso-Mateos JL, Saxena S. Reducing the global burden of depression: Population-level analysis of intervention cost-effectiveness in 14 world regions. Br. J. Psychiatry. 2004;184:393–403.

15. Cuijpers P, van Straten A, Smit F, Mihalopoulos C, Beekman A. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. Am. J. Psychiatry. 2008;165:1272–80.

16. van Zoonen K, Buntrock C, Ebert DD, Smit F, Reynolds CF, Beekman ATF, et al. Preventing the onset of major depressive disorder: A meta-analytic review of psychological interventions. Int. J. Epidemiol. 2014;43:318–29.

17. Bower P, Gilbody S. Stepped care in psychological therapies: Access, effectiveness and efficiency. Narrative literature review. Br. J. Psychiatry. 2005;186:11–7.

18. van Dijk SEM, Pols AD, Adriaanse MC, Bosmans JE, Elders PJM, van Marwijk HWJ, et al. Cost-effectiveness of a stepped-care intervention to prevent major depression in patients with type 2 diabetes mellitus and/or coronary heart disease and subthreshold depression: design of a cluster-randomized controlled trial. BMC Psychiatry. 2013;13:128.

19. Davidson SK, Harris MG, Dowrick CF, Wachtler CA, Pirkis J, Gunn JM. Mental health interventions and future major depression among primary care patients with subthreshold depression. J. Affect. Disord.; 2015;177:65–73.

20. Pols AD, Van Dijk SE, Bosmans JE, Hoekstra T, Van Marwijk HWJ, Van Tulder MW, et al. Effectiveness of a stepped-care intervention to prevent major depression in patients with type 2 diabetes mellitus and/or coronary heart disease and subthreshold depression: A pragmatic cluster randomized controlled trial. PLoS One. 2017;12.

21. Bot M, Pouwer F, Ormel J, Slaets JPJ, de Jonge P. Predictors of incident major depression in diabetic outpatients with subthreshold depression. Diabet. Med. 2010;27:1295–301.

22. Pibernik-Okanovic M, Begic D, Peros K, Szabo S, Metelko Z. Psychosocial factors contributing to persistent depressive symptoms in type 2 diabetic patients: a Croatian survey from the European Depression in Diabetes Research Consortium. J. Diabetes Complications. 2008;22:246–53.

23. Badawi G, Pagé V, Smith KJ, Gariépy G, Malla A, Wang J, et al. Self-rated health: A predictor for the three year incidence of major depression in individuals with Type II diabetes. J. Affect. Disord. 2013;145:100–5.

24. Fisher L, Skaff MM, Mullan JT, Arean P, Glasgow R, Masharani U. A longitudinal study of affective and anxiety disorders, depressive affect and diabetes distress in adults with type 2 diabetes. Diabet. Med. 2008;25:1096–101.

25. Katon W, Ph JR, M EHBLMD, D SRHM, M PCMD, Evette J Ludman Ph, et al. Depression and Diabetes: Factors Associated With Major Depression at Five-Year Follow-Up. Psychosomatics. 2011;50:570–9.

26. Bot M, Pouwer F, Ormel J, Slaets JPJ, de Jonge P. Predictors of incident major depression in diabetic outpatients with subthreshold depression. Diabet. Med. England; 2010;27:1295–301.

27. Nefs G, Pouwer F, Denollet J, Pop V. The course of depressive symptoms in primary care patients with type 2 diabetes: Results from the Diabetes, Depression, Type D Personality Zuidoost-Brabant (DiaDDZoB) Study. Diabetologia. 2012;55:608–16.

28. Doyle F, McGee H, Delaney M, Motterlini N, Conroy R. Depressive vulnerabilities predict depression status and trajectories of depression over 1 year in persons with acute coronary syndrome. Gen. Hosp. Psychiatry.; 2011;33:224–31.

29. Spijkerman TA, Van Den Brink RHS, Jansen JHC, Crijns HJGM, Ormel J. Who is at risk of post-MI depressive symptoms? J. Psychosom. Res. 2005;58:425–32.

30. Pedersen SS, Denollet J, van Gestel YRBM, Serruys PW, van Domburg RT. Clustering of psychosocial risk factors enhances the risk of depressive symptoms 12-months post percutaneous coronary intervention. Eur. J. Cardiovasc.

Prev. Rehabil. 2008;15:203–9.

31. Ossola P, Paglia F, Pelosi A, De Panfilis C, Conte G, Tonna M, et al. Risk factors for incident depression in patients at first acute coronary syndrome. Psychiatry Res. 2015;228:448–53.

32. Kang H, Stewart R, Bae K, Kim S, Shin I, Hong Y, et al. Predictors of depressive disorder following acute coronary syndrome: Results from K-DEPACS and EsDEPACS. J. Affect. Disord.; 2015;181:1–8.

33. Kroenke K SR. The PHQ-9: a new depression diagnostic and severity measure. Psychiatr Ann. 2002;32:509–515.

34. Lamers F, Jonkers CCM, Bosma H, Penninx BWJH, Knottnerus JA, van Eijk JTM. Summed score of the Patient Health Questionnaire-9 was a reliable and valid method for depression screening in chronically ill elderly patients. J. Clin. Epidemiol. 2008;61:679–87.

35. Sheehan DV, Lecrubier Y SK. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatr. 1998;59:22–33.

36. Van Vliet IM, De Beurs E. Het Mini Internationaal Neuropsychiatrisch Interview (MINI): Een kort gestructureerd diagnostisch psychiatrisch interview voor DSM-IV-en ICD-10-stoornissen. Tijdschr. Psychiatr. 2007;49:393–7.

37. Meader N, Mitchell AJ, Chew-graham C, Goldberg D, Rizzo M, Bird V, et al. Case identification of depression in patients with chronic physical health problems: a diagnostic accuracy meta-analysis of 113 studies. 2011;808–20.

38. Moriarty AS, Gilbody S, McMillan D, Manea L. Screening and case finding for major depressive disorder using the Patient Health Questionnaire (PHQ-9): A meta-analysis. Gen. Hosp. Psychiatry; 2015;37:567–76.

39. van der Zwaan GL, van Dijk SEM, Adriaanse MC, van Marwijk HWJ, van Tulder MW, Pols AD, et al. Diagnostic accuracy of the Patient Health Questionnaire-9 for assessment of depression in type II diabetes mellitus and/or coronary heart disease in primary care. J. Affect. Disord. 2015;190:68–74.

40. Pols A, Schipper K, Overkamp D, van Marwijk H, van Tulder M, Adriaanse M. Patients' and practice nurses' perceptions of depression in patients with diabetes type 2 and/or coronary heart disease screened for

subthreshold depression. submitted.

41. Spinhoven P, Ormel J, Sloekers PP, Kempen GI, Speckens AE, Van Hemert AM. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. Psychol Med. 1997;27:363–70.

42. Robins LN, Helzer JE, Croughan J, Ratcliff KS. National Institute of Mental Health Diagnostic Interview Schedule. Arch Gen Psychiatry. 1981;38:381–9.

43. Kriegsman DMW, Penninx BWJH, Van Eijk JTM, Boeke a. JP, Deeg DJH. Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly. A study on the accuracy of patients' self-reports and on determinants of inaccuracy. J. Clin. Epidemiol. 1996;49:1407–17.

44. Twisk J. Different Methods to Analyse the Results of a Randomized Controlled Trial with More Than One Follow-Up Measurement. In: K. van Montfoort, J. Oud & WG, editor. Dev. Stat. Eval. Clin. Trials.; 2014. p. 177–93.

45. McCulloch CE NJ. Generalized Linear Mixed Models. Encyclopedia of Biostatistics. John Wiley & Sons; 2005.

46. Twisk JW. Applied longitudinal data analysis for epidemiology: a practical guide. Cambridge University Press; 2013.

47. Ali S, Stone MA, Peters JL, Davies MJ, Khunti K. The prevalence of co-morbid depression in adults with Type 2 diabetes: a systematic review and meta-analysis. Diabet. Med. 2006;23:1165–73.

48. Comijs HC, Nieuwesteeg J, Kok R, van Marwijk HW, van der Mast RC, Naarding P, et al. The two-year course of late-life depression; results from the Netherlands study of depression in older persons. BMC Psychiatry.; 2015;15:1.

49. de Boer MR, Waterlander WE, Kuijper L, Steenhuis I, Twisk J. Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate. Int. J. Behav. Nutr. Phys. Act. 2015;12:4.

50. Van Buuren, S; Groothuis-oudshoorn K. mice : Multivariate Imputation by Chained. J. Stat. Softw. 2011;45.

51. Rubin DB. Multiple imputation for nonresponse in surveys. John Wiley & Sons; 2004.

52. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. Eur. Urol. 2015;67:1142–51.

53. Heymans MW, Buuren S Van, Knol DL, Van W, Vet HCW De. Variable selection under multiple imputation using the bootstrap in a prognostic study. 2007;10:1–10.

54. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. BMC Med. Res. Methodol. 2009;9:57.

55. Reddy P, Philpot B, Ford D, Dunbar JA. Identification of depression in diabetes: the efficacy of PHQ-9 and HADS-D. Br. J. Gen. Pract. England; 2010;60:e239-45.

56. Cuijpers P, Koole SL, van Dijke A, Roca M, Li J, Reynolds CF. Psychotherapy for subclinical depression: meta-analysis. Br. J. psychiatry. 2014;205:268–74.

57. van't Veer-Tazelaar PJ, van Marwijk HWJ, van Oppen P, van Hout HPJ, van der Horst HE, Cuijpers P, et al. Stepped-care prevention of anxiety and depression in late life: a randomized controlled trial. Arch. Gen. Psychiatry. 2009;66:297–304.

58. Depressie N. M44 NHG-Standaard Depressie. Huisarts&Wetenschap. 2012;55:252–9.

59. Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: A meta-analysis. Cmaj. 2008;178:997–1003.

60. Moons K, Royston P, Vergouwe Y, Grobbee D, Altman D. Prognosis and prognostic research: what, why, and how? BMJ. 2009;338.

61. Kendler KS, Karkowski LM, Prescott CA. Stressful Life Events and Major Depression: Risk Period, Long-Term Contextual Threat, and Diagnostic Specificity. J. Nerv. Ment. Dis. 1998;186:661–9.

62. Assari S, Lankarani MM. Stressful life events and risk of depression 25 years later: Race and gender differences. Front. Public Heal. 2016;4:49.

63. Cuijpers P, Smit F, Willemse G. Predicting the onset of major depression in subjects with subthreshold depression in primary care: A prospective study. Acta Psychiatr. Scand. 2005;111:133–8.

64. Vink D, Aartsen MJ, Schoevers RA. Risk factors for anxiety and depression in the elderly: A review. J. Affect. Disord. 2008;106:29–44.

65. Steyerberg E, Moons KGM, van der Windt D, Hayden J, Perel P, Schroter S, et al. Prognosis research strategy PROGRESS series 3: prognostic model research. PLoS Med. 2013;10:e1001381.

Table 1 Patients' baseline characteristics at baseline in intervention group, care as usual group and total sample

Characteristics	Total sample (N=236)	Intervention (N=96)	Care as usual (N=140)
Female	107/236 (45.3)	42/96 (43.8)	65/140 (46.4)
Age, mean (SD)	67.5 (10.0)	67.8 (9.2)	67.3 (10.5)
Stressful life-event	112/210 (53.3)	48/89 (53.9)	64/121 (52.9)
Positive history of depression	113/210 (53.8)	54/89 (60.7)	59/121 (48.8)
ICPC diagnosis DM2 and/or CHD			
Diabetes Mellitus type 2 (DM2)	88/236 (37.3)	38/96 (39.6)	50/140 (35.7)
Coronary Heart Disease (CHD)	86/236 (36.4)	36/96 (37.5)	50/140 (35.7)
DM2 and CHD	62/236 (26.3)	22/96 (22.9)	40/140 (28.6)
More than 3 chronic diseases	98/210 (46.7)	38/89 (42.7)	60/121 (49.6)
PHQ-9 at baseline, mean (SD)	9.4 (3.2)	9.5 (3.1)	9.3 (3.2)
Anxiety HADS, mean (SD)	6.5 (3.8)	6.9 (3.7)	6.3 (3.9)
Depression HADS, mean (SD)	6.5 (3.8)	6.9 (3.9)	6.1 (3.7)
Marital status			
Married/living together	122/220 (55.5)	55 (61.1)	67/130 (51.5)
Single/divorced/widowed	98/220 (44.5)	35 (38.9)	63/130 (48.5)
Both parents born in the Netherlands	186/220 (84.5)	74/90 (82.2)	112/130 (86.2)
Rural residential area	99/236 (41.9)	42 (43.8)	57/140 (40.7)
Unemployed/sick	26/220 (11.8)	12/90 (13.3)	14/130 (10.8)
Level of education			
Low	89/220 (40.5)	33/90 (36.7)	56/130 (43.1)
Average	60/220 (27.3)	22/90 (24.4)	38/130 (29.2)
High	71/220 (32.3)	35/90 (38.9)	36/130 (27.7)
Current smoker	39/219 (17.8)	16/90 (17.8)	23/129 (17.8)
Alcohol use above norm	63/219 (28.8)	29/90 (32.2)	34/129(26.4)
Exercise under norm	141/219 (64.4)	56/90 (62.2)	85/129 (65.9)
BMI, mean (SD)	28.9 (6.1)	29.4 (6.8)	28.5 (5.6)
Locus of Control, mean (SD)	7.9 (4.2)	8.3 (4.2)	7.6 (4.1)
Social support, mean (SD)	36.3 (9.2)	35.8 (9.0)	36.7 (9.5)
Dysthymia	13/236 (5.5)	6/96 (6.3)	7/140 (5.0)
Onset of depression after age of 55	101/210 (48.1)	38/89 (42.7)	63/121 (52.1)

Numbers are percentages unless stated otherwise; Abbreviations: BMI = Body Mass Index; PHQ-9, Patient Health Questionnaire-9; HADS, Hospital Anxiety and Depression Scale; SD, Standard Deviation.

Table 2 Results of the mixed model and GEE long-term effectiveness analyses

Cumulative incidence of depression (n/N) %	Intervention	Care as usual	Corrected analyses*		Crude analyses	
Baseline	0	0	**OR (95%CI)**	**P-value**	**OR (95%CI)**	**P-value**
T6	(5/84) 6.0	(10/125) 8.0	0.82 (0.19; 3.51)	0.79	0.90 (0.32; 2.50)	0.84
T12	(9/82) 11.0	(12/118) 10.2	1.44 (0.46; 4.47)	0.53	1.20 (0.49; 2.92)	0.70
T24	(13/77) 16.9	(17/105) 16.2	1.23 (0.50; 3.02)	0.66	1.11 (0.51; 2.44)	0.79
Overall effect	n.a	n.a	1.37 (0.52;3.55)	0.52	1.11 (0.49;2.49)	0.80
PHQ mean (SD)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	9.53 (3.14)	9.28 (3.23)	**B (95%CI)**	**P-value**	**B (95%CI)**	**P-value**
T3	6.68 (4.55)	6.58 (4.21)	-0.39 (-1.52; 0.74)	0.50	-0.03 (-1.17; 1.11)	0.96
T6	6.10 (4.43)	6.12 (4.41)	-0.37 (-1.50; 0.76)	0.52	-0.17 (-1.30; 0.95)	0.76
T9	6.28 (4.31)	6.46 (4.51)	-0.48 (-1.62; 0.65)	0.40	-0.40 (-1.53; 0.73)	0.49
T12	6.60 (5.23)	6.29 (4.46)	-0.09 (-1.20; 1.02)	0.88	-0.03 (-1.13; 1.07)	0.96
T24	5.81 (4.76)	5.15 (4.33)	0.00 (-1.18; 1.19)	0.88	0.02 (-1.15; 1.19)	0.97
Overall effect	n.a	n.a	0.29 (-1.15; 0.58)	0.52	-0.13 (-0.99; 0.73)	0.77
Perceived recovery (%)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	n.a	n.a	**OR (95%CI)**	**P-value**	**OR (95%CI)**	**P-value**
T3	40.3%	49.5%	0.78 (0.42; 1.45)	0.44	0.64 (0.36; 1.15)	0.14
T6	48.8%	45.5%	1.46 (0.79; 2.69)	0.23	1.15 (0.65; 2.02)	0.64
T9	55.0%	48.7%	1.47 (0.79; 2.75)	0.22	1.30 (0.74; 2.30)	0.91
T12	55.6%	58.1%	1.04 (0.56; 1.92)	0.91	0.91 (0.51; 1.61)	0.74
T24	68.0%	57.1%	2.38 (1.21; 4.67)	0.01	2.04 (1.08; 3.87)	0.03
Overall effect	n.a	n.a	1.32 (0.87; 2.00)	0.19	1.10 (0.75; 1.62)	0.61
HADS-A mean (SD)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	6.91 (3.74)	6.25 (3.90)	**B (95%CI)**	**P-value**	**B (95%CI)**	**P-value**
T3	6.35 (4.04)	6.29 (3.97)	-0.27 (-1.13; 0.60)	0.54	-0.13 (-1.00; 0.74)	0.76
T6	5.70 (4.10)	6.63 (4.00)	-1.04 (-1.91; -0.18)	0.02	-1.04 (-1.91; -0.18)	0.02
T9	6.16 (4.24)	6.03 (4.04)	-0.49 (-1.35; 0.38)	0.27	-0.45 (-1.31; 0.42)	0.31
T12	5.77 (4.69)	5.83 (3.99)	-0.50 (-1.37; 0.38)	0.27	-0.43 (-1.31; 0.44)	0.33
T24	5.45 (4.46)	5.06 (3.90)	-0.59 (-1.50; 0.31)	0.20	-0.48 (-1.38; 0.43)	0.30
Overall effect	n.a	n.a	-0.59 (-1.23; 0.06)	0.08	-0.52 (-1.17; 0.13)	0.12
HADS-D mean (SD)	**Intervention**	**Care as usual**	**Corrected analyses***		**Crude analyses**	
Baseline	6.93 (3.87)	6.11 (3.73)	**B (95%CI)**	**P-value**	**B (95%CI)**	**P-value**
T3	6.14 (4.16)	6.21 (3.87)	-0.26 (-1.12; 0.60)	0.55	-0.29 (-1.15; 0.56)	0.51
T6	5.82 (3.79)	5.75 (4.03)	-0.22 (-1.07; 0.64)	0.62	-0.32 (-1.18; 0.53)	0.46
T9	6.36 (4.04)	6.07 (4.08)	-0.21 (-1.06; 0.65)	0.63	-0.24 (-1.09; 0.61)	0.58
T12	6.09 (4.20)	6.11 (4.22)	-0.41 (-1.27; 0.46)	0.36	-0.50 (-1.36; 0.36)	0.26
T24	5.59 (4.66)	4.92 (3.90)	-0.41 (-1.30; 0.48)	0.37	-0.48 (-1.37; 0.41)	0.29
Overall effect	n.a	n.a	-0.30 (-0.94; 0.33)	0.35	-0.37 (-1.00; 0.26)	0.25

Abbreviations: 95%CI, 95% Confidence Interval; HADS-A, Hospital Anxiety and Depression Scale-Anxiety; HADS-D, Hospital Anxiety and Depression Scale-Depression; n.a, not applicable; PHQ-9, Patient Health Questionnaire-9;

*Corrected for: baseline values of the outcome, age, gender, marital status, employment status, level of education, co-existence of DM2 and CHD, alcohol use, number of depressive episodes in history and age of onset of depression. The baseline value of the outcome is not added as an extra variable in the corrected analyses of the overall effects since it is already incorporated in the crude overall analyses.

Table 3 Multivariable prediction model of incident depression during two-year follow-up

Predictor	RC	OR	95% CI	P-value
Female sex		-	-	-
Age		-	-	-
Somatic disorder DM2 CHD DM2and CHD		-	-	-
History of depression		-	-	-
Baseline depression scores	0.32 p.p.i.	1.37	1.20; 1.55	0.00
Baseline anxiety scores	0.12 p.p.i.	1.13	1.02; 1.25	0.01
Stressful life-event in past year	0.74	2.10	1.02; 4.32	0.04
>3 chronic illnesses	0.78	2.19	1.12; 4.25	0.02
Randomization status I vs C	0.14	1.15	0.58; 2.29	0.68

RC regression coefficient; p.p.i. per point increase; 95% CI 95% confidence interval; OR odds ratio, an OR > 1 reflects a higher probability the outcome an incident depression and an OR < 1 reflects a lower probability compared with the reference category. OR estimated after multiple imputation (n = 25 datasets) with p-value of 0.157. Linear predictor corrected after bootstrapping = -4.1147 + 0.131* Randomization status + 0.7167* >3 chronic illnesses + 0.680* stressful life-event in past year + 0.1118* baseline anxiety scores + 0.2868* baseline depression scores

SUPPLEMENTARY INFORMATION

Contributors

AP constructed the design of this study, performed all statistical analyses and drafted the manuscript. MCA, MvT, HvM constructed the design of the study and revised the manuscript. JB and SvD constructed the design of the Step-Dep study and revised the manuscript. MH collaborated on the statistical analyses and revised the manuscript. The final manuscript was read and approved by all authors.

Funding

This study is funded by ZonMw, the Netherlands Organisation for Health Research and Development (project number 80-82310-97-12110). The sponsor had no role in the design and conduct of the present study or in the writing of the manuscript.

Competing Interests

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Trial registration and ethical approval

The study was performed in accordance with the declaration of Helsinki (2008) and the Dutch Medical Research involving Human Subjects Act (WMO). The protocol was approved by the medical ethics committee of the VU University Medical Centre (NL39261.029.12, registration number 2012/223), and registered in the Dutch Trial Register (registration number 3715).

Data sharing

Full dataset and statistical code is available from the corresponding author. Consent was not obtained but the presented data are anonymised and risk of identification is low.

Transparency

The corresponding author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; no important aspects of this study have been omitted; and any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Acknowledgements

The authors would like to thank Marcella van der Linden, Lucca Vledder and Mieke Schlattmann for their contribution in the data collection for this study and Jos Twisk for his help in the long-term effectiveness analyses. We also would like to thank all the participating general practices and the research networks of general practitioners (ANH, THOON and LEON) for their participation and collaboration in the implementation and execution of the study. Furthermore, this study has been possible thanks to all Step-Dep participants.

TRIPOD

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	4-5
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	5
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	6
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	6
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	6
	5b	D;V	Describe eligibility criteria for participants.	6
	5c	D;V	Give details of treatments received, if relevant.	7
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	6
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	6
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	7
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	6
Sample size	8	D;V	Explain how the study size was arrived at.	Ref protocol
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	8
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	7-8
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	8-9
	10c	V	For validation, describe how the predictions were calculated.	n.a.
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	9
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	n.a.
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	n.a.
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	n.a.
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	10
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	10, table 1
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	n.a.
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	n.a.
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	n.a.
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	Table 3
	15b	D	Explain how to the use the prediction model.	10-11
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	10-11
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	n.a.
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	12
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	n.a.
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	12-13
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	13-14
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	24-26
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	24

TRIPOD Checklist: Prediction Model Development and Validation

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.